

# Best Practices for Using Geographic Data in Transit: A Location Referencing Guidebook

Defining Geographic Locations of Bus Stops, Routes and other  
Map Data for ITS, GIS and Operational Efficiencies

April 2005

Available from:

Federal Transit Administration  
U.S. Department of Transportation  
Washington, DC 20590

Report Number: FTA-NJ-26-7044-2003.1

**Best Practices for Using Geographic Data in Transit:**  
**A Location Referencing Guidebook**  
Defining Geographic Locations of Bus Stops, Routes and  
Other Map Data for ITS, GIS and Operational Efficiencies

**TABLE OF CONTENTS**

<b>FOREWORD .....</b>	<b>1</b>
<b>ACKNOWLEDGEMENTS .....</b>	<b>2</b>
<b>EXECUTIVE SUMMARY .....</b>	<b>4</b>
<b>SECTION 1 INTRODUCTION.....</b>	<b>10</b>
1.1 Barriers to Sharing and Using Spatial Data.....	10
1.2 Impacts of Barriers .....	11
1.3 Definition of Location Referencing .....	11
1.4 Why a Location Referencing Best Practices Guidebook?.....	12
1.5 Guidebook Development Process .....	12
1.6 Goals and Objectives of the Location Referencing Guidebook .....	13
1.7 Scope of Location Referencing Guidebook.....	13
1.7.1 Different Audiences .....	13
1.7.2 Transit Feature Data Emphasis .....	14
1.7.3 Different Perspectives .....	14
<b>SECTION 2 ROADMAP FOR MANAGERS .....</b>	<b>16</b>
2.1 Section Overview.....	16
2.1.1 Purpose.....	16
2.1.2 Problem .....	16
2.2 Role of Transit Management .....	17
2.3 Roadmap for Managers.....	17
2.3.1 Ensure Business Goals Drive Technology Investments .....	18
2.3.2 Create an Enterprise-wide Approach .....	19
2.3.3 Foster Consideration of Regional Partners .....	20
2.3.4 Align Planning Efforts .....	20

2.3.5	Budget to Protect Investments .....	21
2.3.6	Increase the Effectiveness of Technology Procurements .....	21
2.3.7	Require Education on Technology Options and Invest in Training .....	21
2.3.8	Influence Strategic Technology Decisions and Implementing Standards .....	22
2.3.9	Provide Ongoing Oversight and Support.....	23
<b>2.4</b>	<b>Think Beyond .....</b>	<b>23</b>
<b>SECTION 3 CREATING AN EFFECTIVE SPATIAL INFORMATION INFRASTRUCTURE .....</b>		<b>24</b>
<b>3.1</b>	<b>Section Overview.....</b>	<b>24</b>
3.1.1	Problem.....	24
3.1.2	Purpose.....	25
<b>3.2</b>	<b>Develop Infrastructure with a Shared Vision and Goals .....</b>	<b>25</b>
3.2.1	An Integrated Business Enterprise.....	25
3.2.2	A Strategic Information Framework.....	26
<b>3.3</b>	<b>Plan for Growth .....</b>	<b>27</b>
<b>3.4</b>	<b>Managing Spatial Relationships: Core Data and a Data Model.....</b>	<b>28</b>
3.4.1	Identify and Develop Core Data .....	28
3.4.2	Building a Data Model.....	30
<b>3.5</b>	<b>Develop a Common Graphical User Interface for Spatial Data.....</b>	<b>30</b>
<b>3.6</b>	<b>Implement Standards .....</b>	<b>31</b>
<b>3.7</b>	<b>Use Effective Database Access Methods .....</b>	<b>31</b>
<b>3.8</b>	<b>Maintain Spatial Data Quality.....</b>	<b>32</b>
<b>3.9</b>	<b>Improve Project Management and Life Cycle Approach .....</b>	<b>32</b>
<b>3.10</b>	<b>Resources .....</b>	<b>33</b>
<b>SECTION 4 GENERAL PROCUREMENT GUIDELINES.....</b>		<b>35</b>
<b>4.1</b>	<b>Overview .....</b>	<b>35</b>
4.1.1	Problem.....	35
4.1.2	Purpose.....	35
<b>4.2</b>	<b>FTA’s National ITS Architecture Policy on Transit Projects .....</b>	<b>36</b>
<b>4.3</b>	<b>Collaboration Among Internal and External Organizations.....</b>	<b>36</b>
<b>4.4</b>	<b>Intellectual Property and Data Sharing Issues .....</b>	<b>37</b>

<b>4.5</b>	<b>Procurement Provisions for Data and Interfaces .....</b>	<b>39</b>
<b>4.6</b>	<b>Standards and Internal Data Policies .....</b>	<b>39</b>
4.6.1	The Open-Source Movement in Government.....	40
<b>4.7</b>	<b>Evaluating Proposals .....</b>	<b>41</b>
<b>4.8</b>	<b>Resources .....</b>	<b>41</b>
<b>SECTION 5</b>	<b>THE BASE MAP .....</b>	<b>42</b>
<b>5.1</b>	<b>Overview .....</b>	<b>42</b>
5.1.1	Need for a Base Map.....	42
5.1.2	Industry Pressures and Agency Inefficiencies .....	42
5.1.3	Problems in Managing a Transit Base Map.....	43
5.1.4	Purpose.....	43
<b>5.2</b>	<b>Definitions.....</b>	<b>44</b>
<b>5.3</b>	<b>Related Transit and ITS Applications .....</b>	<b>45</b>
<b>5.4</b>	<b>Recommended Practices.....</b>	<b>45</b>
5.4.1	Single Base Map .....	45
5.4.2	Define Stakeholders and Their Requirements .....	46
5.4.3	Essential Data Elements for a Transit Base Map.....	46
5.4.4	The Decision to Build or Buy a Transit Base Map.....	47
5.4.5	Lessons Learned by Transit Agencies in Using Commercial Base Maps .....	50
5.4.6	Guidelines for Evaluating Commercial Base Maps.....	51
5.4.7	Guidelines for Evaluating Commercial Base Map Providers .....	53
<b>5.5</b>	<b>Technical Discussion.....</b>	<b>54</b>
5.5.1	Lesson Learned: Need for Unique Segment IDs and Links to Historical Identifiers.....	54
5.5.2	Patches .....	55
5.5.3	Hidden Costs of Proprietary Data Formats.....	55
<b>5.6</b>	<b>Think Beyond .....</b>	<b>56</b>
<b>SECTION 6</b>	<b>POPULATING AND ORGANIZING THE BASE MAP WITH TRANSIT FEATURES 57</b>	
<b>6.1</b>	<b>Overview .....</b>	<b>57</b>
6.1.1	Problem.....	57
6.1.2	Purpose.....	59
<b>6.2</b>	<b>Location Referencing Methods Defined.....</b>	<b>59</b>
<b>6.3</b>	<b>Transit Feature Definitions and Attributes.....</b>	<b>60</b>

6.3.1	Point Features.....	62
6.3.2	Line and Path Features.....	64
6.3.3	Polygon Features.....	66
6.3.4	Relationships Among Transit Features.....	68
6.3.5	Associating Performance Measures with Transit Features.....	71
<b>6.4</b>	<b>Recommended Practices for Transit Features .....</b>	<b>71</b>
6.4.1	Identifying Transit Feature Building Blocks and Associating Them With the Base Map	72
6.4.2	Describing Transit Features .....	74
6.4.3	Building a Data Model.....	77
6.4.4	Considerations in Managing Location Referencing Methods .....	77
6.4.5	Building a Location Table .....	78
6.4.6	Storing Transit Features.....	80
<b>6.5</b>	<b>Resources .....</b>	<b>81</b>
<b>SECTION 7</b>	<b>SPATIAL DATA QUALITY .....</b>	<b>83</b>
<b>7.1</b>	<b>Overview .....</b>	<b>83</b>
7.1.1	Problem.....	84
7.1.2	Purpose.....	84
<b>7.2</b>	<b>What is Quality? .....</b>	<b>84</b>
7.2.1	Data Quality Principles and Data Management.....	85
7.2.2	Data Standards Help .....	85
7.2.3	Data Quality Elements and Determining Fitness for Use.....	85
<b>7.3</b>	<b>Data Quality Elements.....</b>	<b>86</b>
7.3.1	Positional Accuracy .....	87
7.3.2	Attribute Accuracy.....	87
7.3.3	Logical Consistency.....	88
7.3.4	Completeness .....	89
7.3.5	Lineage.....	89
7.3.6	Data Quality Elements and Subcomponents.....	90
<b>7.4</b>	<b>Recommended Practices.....</b>	<b>91</b>
7.4.1	Defining a Quality Process .....	91
7.4.2	Internal Data Policies and Standards .....	92
7.4.3	Enterprise IT and ITS Standards.....	92
7.4.4	Data Ownership .....	93
7.4.5	Data Committee and Organizational Data Surveys .....	93
7.4.6	Metadata.....	93
7.4.7	Configuration Management .....	94
7.4.8	Quality Checks.....	95
<b>7.5</b>	<b>Resources .....</b>	<b>97</b>

## **SECTION 8 ITS APPLICATIONS: POLICY REQUIREMENTS AND SPATIAL ISSUES 99**

<b>8.1</b>	<b>Section Overview.....</b>	<b>99</b>
8.1.1	Problem.....	99
8.1.2	Purpose.....	99
<b>8.2</b>	<b>Spatial Representation Issues Within a Systems Engineering Context.....</b>	<b>100</b>
8.2.1	Planning and Requirements Analysis Phases.....	101
8.2.2	Alternative Analysis Phase.....	103
8.2.3	Procurement Phase.....	105
8.2.4	Design Phase.....	106
8.2.5	Deployment Phase.....	106
8.2.6	Operations and Maintenance Phase.....	107
<b>8.3</b>	<b>Selected ITS Application Issues and Comments.....</b>	<b>107</b>
8.3.1	Bus Stop Inventory.....	107
8.3.2	On-Board Systems.....	110
8.3.3	Customer Information System (Trip Itinerary Planning).....	112
8.3.4	Real-time Customer Information.....	112
8.3.5	Ridematching Software.....	113
8.3.6	Transit Signal Priority.....	113
8.3.7	Paratransit.....	113
<b>8.4</b>	<b>Resources.....</b>	<b>114</b>

## **SECTION 9 SHARING BETWEEN APPLICATIONS: COMMON INTERFACES AND MAINTENANCE SYNCHRONIZATION..... 115**

<b>9.1</b>	<b>Overview.....</b>	<b>115</b>
9.1.1	Problem.....	115
9.1.2	Purpose.....	116
<b>9.2</b>	<b>General Requirements That Support Interfaces and Location Referencing.....</b>	<b>117</b>
<b>9.3</b>	<b>Synchronizing Dataset Maintenance Cycles.....</b>	<b>119</b>
9.3.1	Base Map Update Process.....	121
9.3.2	Building a Base Map Data Configuration Managmeent Plan.....	123
<b>9.4</b>	<b>Impact of Client/Server Models.....</b>	<b>124</b>
9.4.1	Client-Server Approach and Issues.....	124
9.4.2	Alternative Approach.....	124
9.4.3	Middleware and Web Services.....	126
<b>9.5</b>	<b>Application Integration Metadata.....</b>	<b>127</b>
<b>9.6</b>	<b>Resources.....</b>	<b>129</b>

## **SECTION 10 EXTERNAL AND INTERNAL AGENCY COORDINATION..... 130**

<b>10.1 Overview .....</b>	<b>130</b>
10.1.1 Problem .....	130
10.1.2 Purpose.....	131
<b>10.2 Coordinating with External Agencies .....</b>	<b>131</b>
10.2.1 External Data Sharing Approaches .....	131
10.2.2 Transit Agency Examples of Regional Data Sharing .....	132
10.2.3 Motivations for Regional Data Sharing Efforts .....	133
10.2.4 Important Data Partners for a Transit Agencies .....	134
10.2.5 Impacts on Transit Agencies.....	135
10.2.6 Critical Success Factors .....	136
<b>10.3 Coordinating Internally Within the Agency.....</b>	<b>137</b>
10.3.1 Spatial Data Sharing Internally Among Organizational Units .....	137
10.3.2 Stakeholders .....	138
10.3.3 Procedural and Technical Considerations.....	138
10.3.4 Recommendations.....	139
<b>10.4 Resources .....</b>	<b>140</b>

## **SECTION 11 THE IMPORTANCE AND ROLE OF METADATA IN EFFECTIVE SPATIAL DATA USAGE ..... 141**

<b>11.1 Section Overview.....</b>	<b>141</b>
11.1.1 Definition of Metadata .....	141
11.1.2 Problem .....	142
11.1.3 Purpose.....	142
<b>11.2 Benefits of Metadata .....</b>	<b>142</b>
11.2.1 Benefits .....	142
11.2.2 Why Should Transit Management Care About Metadata? .....	143
<b>11.3 Standards for Transit Metadata.....</b>	<b>145</b>
<b>11.4 What Information Belongs in a Metadata Record?.....</b>	<b>146</b>
11.4.1 Metadata Record Checklist .....	146
11.4.2 Tools for Metadata Management .....	147
11.4.3 Priorities for Implementing Metadata for the First Time.....	148
11.4.4 Finding and Identifying Historical Data .....	148
<b>11.5 Transit Agency Examples .....</b>	<b>149</b>
<b>11.6 Think Beyond .....</b>	<b>150</b>
11.6.1 Quality and Procedures for Measuring Accuracy .....	150
11.6.2 Integration Metadata .....	151

<b>11.7 Resources and References .....</b>	<b>151</b>
<b>APPENDIX A GUIDEBOOK DEVELOPMENT PROCESS .....</b>	<b>A-1</b>
<b>APPENDIX B RECOMMENDATIONS FOR BASE MAP ATTRIBUTES .....</b>	<b>B-1</b>
<b>APPENDIX C TCIP LOCATION REFERENCING METHODS .....</b>	<b>C-1</b>
<b>APPENDIX D THE GEOSPATIAL ONE-STOP'S TRANSIT REFERENCE MODEL</b>	<b>D-1</b>
<b>APPENDIX E ..... ADA COMPLIANCE AND BUS STOP ACCESSIBILITY DECISION TREE.....</b>	<b>E-1</b>
<b>APPENDIX F SAMPLE DATA DICTIONARY FOR BUS STOP SIGN.....</b>	<b>F-1</b>
<b>APPENDIX G TRANSIT APPLICATIONS/FUNCTIONS THAT USE SPATIAL DATA AND SPATIAL ANALYSIS FUNCTIONS (GIS TOOLS) .....</b>	<b>G-1</b>
<b>APPENDIX H CONCEPT OF OPERATIONS FOR TRIP ITINERARY PLANNING .</b>	<b>H-1</b>
<b>APPENDIX I TRI-MET'S CONFIGURATION MANAGEMENT PLAN FOR BASE MAP UPDATE .....</b>	<b>I-1</b>
<b>APPENDIX J ACRONYMS AND GLOSSARY .....</b>	<b>J-1</b>



## FOREWORD

This Guidebook was developed at the request of the transit industry. It provides best practices for both transit managers and technical staff with respect to planning, implementing and using geographic data in transit. The Guidebook discusses issues and best practices for defining and using geographic locations of bus stops, routes and other map data that are needed for successfully implementing Intelligent Transportation Systems (ITS) and Geographic Information Systems (GIS), as well as for obtaining operational efficiencies.

Transit managers play a critical role in the successful use of geographic data in transit. Section 2, titled Roadmap for Managers, includes a high level roadmap comprised of nine steps for managers that will help the operation of transit by improving spatial data usage in analyses, maps, customer information and ITS applications. In addition, the beginning of each section of the Guidebook contains an overview that is targeted to both management and technical staff. The overview orients the reader to purpose of the section with respect to improving transit and the issues to be discussed.

All of the sections of the Guidebook have content that is relevant to Information Technology managers, GIS staff and ITS project managers. In anticipation that sections of the Guidebook may be read and used in a different order than presented, references to other sections are liberally included and the first occurrence of acronyms are spelled out at the start of each section.

### ***NOTICE:***

This document is disseminated under the sponsorship of the United States Department of Transportation in the interest of information exchange. The United States Government assumes no liability for its contents or use thereof.

The United States Government does not endorse manufacturers or products. Trade names appear in the document only because they are essential to the content of the report.

## ACKNOWLEDGEMENTS

This Guidebook is a compilation of the best practices for using geographic data in transit. The Guidebook was made possible through the generous contributions of time and expertise by a diverse and large group of individuals. These individuals are experts in the public sector, consultant sector and vendor sectors of the transit and geospatial industries.

The project was composed of two phases, the Feasibility Study and development of the Guidebook. The Feasibility Study Phase resulted in two Technical Memoranda on transit needs and available standards; the Guidebook development phase resulted in this document. Both efforts were funded under a U.S. Government cooperative agreement FTA C-6, October 1, 1999. This Guidebook benefited from the vision and encouragement of Brian Cronin and his colleagues at the Federal Transit Administration (FTA).

The development of this document is based on voluntary contributions from individuals associated with various public agencies and private companies. Collectively, these participants donated literally hundreds of hours to describe best practices needed by the transit industry, and to review this document. Their time and effort are greatly appreciated. Bibiana Kamler and Jeff Orton serve as co-chairs to this group of dedicated individuals. Tom Sullivan is chair of the Transit Standards Consortium (TSC) Technical Working Group that oversaw this effort. The volunteers include:

Abed Abukar, DART (Dallas)  
Alex Guzman, Miami-Dade Transit  
Anna Flintoft, RTA (Chicago)  
Art Scanlon, Clever Devices  
Benjamin Pecheux, Mitretek  
Bert Jakubs, Tele Atlas  
Bibiana Kamler, Tri-Met (Portland, OR)  
Bill Demarko, NavTech  
Brad Thompson, RTA (Chicago)  
Brendon Ford, Fairfax County (Virginia)  
Brendon Hemily, independent consultant  
Carl Reed, OGC  
Darryl Riley, King County Metro  
Dave Matta, Avail Technology  
David Beal, Bi-State  
David Crout, Tri-Met  
Dick Bolend, Portland METRO Regional Council of Governments  
Don Cooke, GDT  
Emilio Escudero, MTC (Oakland)  
Ernie Ott, ESRI  
Fabian Cevallos, CUTR  
Glenn Newmann, New Jersey Transit  
George Wagner, Siemens  
Georgina (Gina) Del Valle, Miami-Dade Transit

Jim Davis, NYSDOT  
Jim Kemp, New Jersey Transit  
Jim McBride, WMATA  
John Sutton, Cambridge Systematics  
Jon Lutterman, Tri-Met  
Justin Begley, Broward County Transit  
Karen Watkins, Lockheed Martin  
Kevin Maybe, GDT  
Lou Millan, New Jersey Transit  
Mark Bosworth, Portland METRO Regional Council of Governments  
Michael Berman, King County Metro (KCM)  
Mike Marques, Tele Atlas  
Jeff Orton, Kern County, CA  
Patrick Soto, Los Angeles MTA  
Peter Furth, Northeastern University  
Rex Fisher, Tri-Met  
Ron Bargman, GDT  
Shirley Hsaio, Orange County Transportation Authority (OCTA)  
Somitra Saxena, bd Systems (formerly GISTrans)  
Steve Parrot, Lane Transit District  
Terry Bills, bd Systems  
Tim McHugh, Tri-Met  
Tom Friedman, KCM  
Tom Kowalski, UTA  
Tom Passin, Mitretek  
Tom Vaughan, NYSDOT

The consultants who wrote, incorporated comments and edited this Guidebook include:

Paula Okunieff, Systems and Solutions, Inc.  
Teresa Adams, University of Wisconsin-Madison  
Nancy Neuerburg, N-Squared Associates

For her support and encouragement in the last phase of the project, the authors would like to thank Susan Proper of Mitretek for her thoughtful comments, excellent edits and fresh perspective.

Jim Kemp of the Transit Standards Consortium provided important help with logistics and resources. Lisa Durst provided her expertise to improve the look of the Guidebook and to create a single document from all the sections. Finally, Richard Cox of the Society of Automotive Engineers provided support and administered this project from the beginning.

## EXECUTIVE SUMMARY

### PURPOSE

A major issue faced by transit agencies is how to successfully and cost effectively develop, integrate, use and share spatial data. Spatial data, such as the geographic location of bus stops, routes, transit facilities and the regional street network, are critical for the efficient operation of transit. Many Intelligent Transportation Systems (ITS) applications such as Automated Vehicle Location (AVL), Automated Passenger Counters (APC) and Itinerary Planning systems will not operate correctly without appropriate spatial data. A significant barrier to success is how the location of spatial data is characterized or referenced. Transit agencies throughout the United States have been “learning the hard way” about location referencing problems. They have also been creating some valuable solutions and best practices.

Based on industry input, this Guidebook summarizes technical and organizational issues that influence the effectiveness with which spatial data are managed in a transit agency and provides guidance for avoiding and overcoming the barriers that typically pose problems for transit with respect to spatial data. Specifically, this Guidebook presents a set of “best practices” that minimize location-referencing problems and support effective spatial data usage. Given new tools and best practices information, transit agencies will benefit as spatial databases are built and expanded, as Geographic Information System (GIS) and ITS applications are installed and as ITS applications are maintained and operated.

### GUIDEBOOK DEVELOPMENT

This Guidebook is the product of Phase II of the Transit Standards Consortium (TSC) Location Referencing Guidebook project conducted by the TSC under U.S. Government Cooperative Agreement FTA C-6, October 1, 1999. The Guidebook was drafted with oversight from the Location Referencing Guidebook Working Group.

### AUDIENCE AND ORGANIZATION

This Guidebook includes high level materials for transit managers and more detailed information for Information Technology (IT), ITS and GIS managers and technical staff. It is designed to help both non-technical and technical transit staff, by providing both overview materials and detailed best practices. Specifically, it should help:

- Senior Management (identifies best practices related to oversight and policy level issues; describes agency-wide perspective; highlights barriers and responsibilities)
- Project Managers (identifies best practices related to managing

Table 1: *Best Practices for Using Geographic Data in Transit*: Section Titles

Section 1: Introduction
Section 2: Roadmap for Managers
Section 3: Creating an Effective Spatial Information Infrastructure
Section 4: General Procurement Guidelines
Section 5: The Base Map
Section 6: Populating and Organizing the Base Map with Transit Features
Section 7: Spatial Data Quality
Section 8: ITS Applications: Policy Requirements and Spatial Issues
Section 9: Sharing Between Applications
Section 10 External & Internal Agency Coordination
Section 11: Importance and Role of Metadata

projects that use or produce spatial data, including discussion of systems engineering analyses, procurement, and implementation issues)

- GIS/IT/Data Management Staff (identifies best practices related to location referencing and technical development and maintenance of spatial data)

Table 1 lists the titles for the 11 sections of the Guidebook. Additional detail is included in the nine appendices. The sections are organized to be self-contained units that address specific issues relevant to a particular audience and area. Sections 1 through 4 and parts of Section 10 address management and policy issues, with Section 2 focusing on the role of senior managers. Sections 5 through 11 address technical and process issues. The introduction of each section consists of a high level discussion of the problem and purpose of the topic covered by that section; it captures the essence of the section's discussion in non-technical language.

## **PROCESS**

This Guidebook of best practices was developed in two phases. The process of identifying the best practices involved the following four main data collection efforts, one in Phase 1 and the other three in Phase 2:

- In Phase 1, a feasibility study was conducted that gathered information on transit's spatial data requirements and on available spatial data standards.
- In Phase 2, the remainder of the best practices was gathered and the Guidebook was developed starting with a series of interviews. The interviews were completed with a diverse set of 13 analysts and GIS managers at transit agencies and regional authorities. These interviewees were known to be leaders in the area of spatial data management. They worked for a cross-section of agencies, including large, small, regional and local transit agencies.
- Many issues and best practices were derived from a series of ten teleconference workshops for the transit industry that were hosted by the Transit Standards Consortium (TSC). The agendas for the workshops were developed from the synthesized results of the interviews (see Appendix A). The workshop participants included interested parties and invited subject matter experts from transit agencies, counties, metropolitan planning organizations, transit application vendors, GIS software vendors and commercial map database vendors.
- The fourth source of information for deriving best practices included personal communications as follow-up to the workshops, published literature, and internal documents from transit and county agencies.

This Guidebook is the product of a review by workshop participants, Federal Transit Administration, TSC, TCIP and Geo-spatial One Stop staff. The Acknowledgement page includes a list of individuals and their affiliations from all the organizations that contributed to this Guidebook.

## **BACKGROUND**

This Guidebook contains information and best practices derived from a wide range of sources as described in the Process Section above. The Guidebook illustrates strategies and methods in current use at many organizations including Tri-Met, King County Metro, OCTA, Fairfax County, New Jersey Transit, Miami-Dade Transit, and others as listed in the Acknowledgements

section. The Guidebook reflects trends in standards development both in the United States and internationally. It discusses developments in the ITS, IT, GIS and geospatial industries at the time of the development of this document.

### ***Location Referencing Defined***

A geographic location can be “referenced” in many ways, such as by an address or by a latitude and longitude. Effective use of location referencing is needed to implement many of the tools transit agencies want to use to improve operational efficiency and effectiveness. A technical definition of a location referencing method is the position of an entity relative to other entities or to some external frame of reference (e.g., latitude and longitude relative to the spherical geoid).

Within the context of transit, location referencing may be defined as how a transit feature (e.g., bus stop, route) is associated to a base map, or where it is relative to a known location. This definition implies that transit features contain spatial characteristics, e.g., a route, bus stop and fare zone have spatial characteristics. The same location may be described by a variety of different *location referencing methods*, as shown in the following examples:

- Third stop inbound on Route #39
- Ten feet east on the south side from the corner on South Street intersecting with School Street
- Near-side on the corner of South and School streets [on street, cross street]
- 25 Elm Street, Cambridge, MA 02140 USA
- Route #10, mile 1.2, inbound direction
- <lat, long> WGS ‘84
- <Easting, Northing> MA State Plane
- Link #58673901 on the XYZ digital (vector) base map
- Grid 45-245 on the ABC raster base map

Different ITS, GIS and database applications require different location reference methods. Applications that collect and share transit features and their attributes must be able to access, process and transform location references from one method to another. This is one of the most critical issues related to sharing transit data.

### ***Barriers to Sharing Transit Data***

Effective management and exchange of spatial location information enables more efficient service delivery, better quality services, more accurate customer information and improved integration between modes and service providers. However, a variety of barriers exist that limit or prohibit transit agencies from realizing those benefits. These barriers may be due to technical and/or organizational issues. Examples of common barriers to sharing spatial data and supporting business needs that depend on spatial data are included below.

- Multiple databases and applications can introduce barriers due to
  - Inconsistent naming conventions
  - Different transit data models
  - Embedded GIS-like functionality with different data representations
  - Differences in data source quality
  - Transformations and translations

- Poor documentation (metadata)
- Use of different location referencing systems
- Level of knowledge of standards, existing tools and utilities
- Quality of Request for Proposal (RFP) and systems specifications
- Definition and execution of operational and data maintenance responsibilities

These barriers can occur both within a transit agency and between organizations in a region when data sharing or systems integration is required. Further, the barriers can emerge at the various stages of the data life cycle.

Unfortunately for transit, the barriers can prevent the unambiguous referencing of transit feature location attributes. When locations are not clearly referenced, they cannot be transformed from one location referencing method to another, then data sharing is challenged, and map and analysis errors can be produced.

### ***Impacts of Barriers***

Spatial data problems can be costly for transit agencies. When the implementation of a new ITS or information reporting systems is delayed because of spatial data problems, the costs of the delays can be measured in additional labor expenditures for staff, consultants and/or vendors. Missed opportunity costs also occur from not deploying the systems in a timely manner. Examples of potentially vulnerable ITS applications include Itinerary Planning, AVL, APC and annunciators on the vehicles. Other impacts may include the following:

- Loss of cost savings from eliminating duplicative spatial data maintenance
- Loss of cost savings from eliminating duplicative spatial database development
- Difficulties in conducting needed operational analyses
- Missed opportunities to leverage funds by sharing data regionally
- Loss of management and public confidence due to inconsistent information issued by non-integrated systems.

### ***BEST PRACTICES***

Highlights of the Best Practices may be summarized into three areas:

- Managing the Technology Investment (management perspective)
- Managing the Life-Cycle Maintenance of the Base Map and Associated Transit Feature Data
- Issues related to Sharing Spatial Data among ITS and other Applications

The best practices are summarized below.

#### ***Best Practices for Managing the Technology Investment***

Effective enterprise-wide use of spatial data, and indeed of all information technology, requires a successful integration of traditional management practices with good systems engineering practices. The following nine practices support successful decision-making and operations by improving spatial data usage in analyses, maps, customer information and ITS applications. These practices are highlighted from the perspective of senior transit managers in Section 2,

Roadmap for Managers. Section 3, Creating an Effective Spatial Information Infrastructure, provides additional information on many of the best practices listed below.

1. Ensure Business Goals Drive Technology Investments
2. Create an Enterprise-wide Approach
3. Foster Consideration of Regional Partners
4. Align Planning Efforts
5. Budget to Protect Investments
6. Increase the Effectiveness of Technology Procurements
7. Require Education on Technology Options and Invest in Training
8. Influence Strategic Technology Decisions and Implement Standards
9. Provide Ongoing Oversight and Support

### ***Best Practices for Life-Cycle Maintenance of the Base Map and Associated Transit Feature Data***

The majority of the Guidebook recommendations deal with spatial data life-cycle issues and barriers that can occur in the planning, development, operation and maintenance stages. Recommendations are based on implementing a systems engineering approach to gather requirements from stakeholders, scope projects, assess fitness of existing systems and data, and understand data and application dependencies. The recommendations also advocate ensuring sufficient enterprise data resources (e.g., database administration staff, data owners, budgeting for updates), and enforcing corporate data policies with respect to data descriptions and usage.

Section 4, General Procurement Guidelines, discusses issues that should be addressed in the procurement phase to avoid problems during the data life cycle. Section 4 also addresses issues such as understanding the intellectual property restrictions on the access and use of commercial data or data stored in commercial applications.

The most important best practices include:

1. *Systems Engineering Approach*: Define stakeholders and their requirements using a systems engineering approach.
2. *Single, Enterprise-wide Base Map*: Adopt a single base map to maintain and propagate it to different applications for further customization as needed; evaluate the options for acquiring a base map that meets transit business needs across the agency. The development of a base map is discussed in Sections 5, The Base Map, and Section 10, External and Internal Agency Coordination.
3. *Core Transit Data*: Define a corporate transit feature dictionary and define an enterprise data model; store and manage the transit features centrally; establish procedures for collecting and updating data.
4. *Location Reference Authority*: Develop an “authority” such as a Location Table to support transit spatial feature portability and storing transit feature data. Include several, key location referencing methods in the feature description or in the Location Table.
5. *Data Maintenance*: Develop a data maintenance program and synchronize the maintenance schedules for the various applications that use spatial data from an enterprise-wide perspective.



6. *Data Quality and Metadata:* In the metadata, document the quality of the spatial dataset, the data collection procedures, and the identification and access methods.
7. *Unique Identifiers for Core Transit Feature:* Use unique ID numbers for features; do not recycle or delete ID numbers during the life of a feature inventory.
8. *Addressing Format:* Consider using the same format as the U.S. Postal Service or the Census Bureau. Many commercial GIS packages provide several standardized formats as options. The most important best practice is to be CONSISTENT across all applications within an agency.
9. *Training and skill level:* Hire a Database Administrator (DBA) to manage the core transit data. A person knowledgeable in GIS should manage the base map and associated transit feature data. Invest in training for staff that develop, maintain and use the spatial data.

### ***Best Practices for Sharing Spatial Data among ITS and other Applications***

Effective life-cycle management enables the unambiguous sharing of information between applications. The best practices related to ITS and sharing among applications assumes that a robust data management approach is in place. Key best practices are included below that further facilitates effective spatial data sharing among applications.

1. *Standard Data Access Methods:* Define standard access methods to share transit feature data and other spatial data. Centralize the methods as stored procedures or middleware.
2. *Integration Metadata:* Manage application integration using an enterprise-wide repository of integration metadata. Develop the repository incrementally as each project is developed.
3. *Data Update Synchronization:* Adopt a data maintenance policy that ensures that applications start with the same datasets. Synchronize data update schedules, and standardize procedures.
4. *Manage Data Centrally:* Use a GIS and an enterprise-wide database management system to centralize corporate spatial data.

Issues and guidelines are described for several ITS applications. These short “crib sheets” describe some data and integration issues that should be considered as part of the planning process. The ITS packages described include Bus Stop Inventory, On-board Systems, Paratransit, Customer Information, Real-time Passenger Information and others.

To successfully share information, requires a deliberate degree of standardization across the enterprise. Users must share a common vocabulary including data meaning and format. To benefit fully, the data must be stored and accessed similarly. Centralizing the data and access methods will ensure that if there are changes made, then their impacts will be more easily managed. There are several information technology, ITS and transit standards emerging to support interoperability. The efforts include eXtensible Markup Language (XML) and its family of standards, Transit Communications Interface Profile (TCIP), Geospatial One-Stop (OGC), NTCIP family of standards, and other ITS standards. Because there are always new and emerging methodologies and techniques, the list will definitely grow in the future and benefit transit.

## **SECTION 1 INTRODUCTION**

Much of the data used in planning, operating and assessing the performance of a transit agency have a spatial or geographic component. As a result, spatial data can be a significant component of the information resources of a transit agency. For example, passengers and transit employees need to know where the buses are. Planners want to know where the passengers load and alight; and community relations staff want to be able to show the public where service changes will occur.

Maintenance staff needs to know where work is required. Similarly, individuals in the safety and security units need to know the locations of problem areas. Appendix G lists the many business areas of transit and provides examples of each business area of how spatial data and geographic information system (GIS) functions can support transit.

Given the importance to the transit industry of being able to use spatial data effectively to support its business, the Federal Transit Administration (FTA) funded the development of this Guidebook for the transit industry. The intent of this Guidebook, *Best Practices for Using Geographic Data in Transit: A Location Referencing Guidebook*, is to help transit overcome many of the spatial location referencing issues that are creating inefficiencies and barriers to transit's sharing and use of spatial data.

The location of many transit variables and attributes can be spatially represented or referenced, such as the location of routes and facilities, ridership, accident locations and potential markets. To effectively display transit attributes, the spatial representation of the underlying street network and service area is critical. The spatial data can be used in GIS maps, new analyses, visual representations of findings, providing various types of customer information and in providing the data needed by new Transit Intelligent Transportation Systems (ITS). Many of the ITS applications not only require spatial data to operate, but also create data with spatial components.

### **1.1 Barriers to Sharing and Using Spatial Data**

Effective management and exchange of spatial location information enables more efficient service delivery, better quality services, accurate customer information and improved integration between modes and service providers. However, a variety of barriers exist that limit or prohibit transit agencies from realizing those benefits. These barriers may be due to technical and/or organizational issues. Examples of common barriers to sharing spatial data and supporting business needs that depend on spatial data include:

- Lack of adequate strategic direction and standards
- Multiple databases and applications can introduce barriers due to
  - Inconsistent naming conventions
  - Different transit data models
  - Embedded GIS-like functionality with different data representations
  - Differences in data source quality
  - Transformations and translations
  - Poor documentation (metadata)

- Use of different location referencing systems
- Level of knowledge of standards, existing tools and utilities
- Quality of Request for Proposal (RFP) documents and systems specifications
- Definition and execution of operational and data maintenance responsibilities

These barriers can occur both within a transit agency and between organizations in a region when data sharing or systems integration is required. Further, the barriers can emerge at the various stages of the data life cycle.

Unfortunately for transit, the barriers can prevent the unambiguous referencing of transit feature location attributes. When locations are not clearly referenced, data sharing is challenged and map and analysis errors can be produced.

## **1.2 Impacts of Barriers**

Spatial data problems can be costly for transit agencies. When new ITS or information reporting systems have spatial data problems and are delayed, the costs can be measured in additional staff, consultant or vendor labor expenditures. Missed opportunity costs also occur from not deploying the systems in a timely manner. Examples of potentially vulnerable ITS applications include Itinerary Planning, Automated Vehicle Location (AVL), Automated Passenger Counting (APC) and on-bus annunciators. Other impacts may include the following:

- High cost of duplicative spatial data maintenance
- Loss of cost savings from eliminating duplicative spatial database development
- Difficulties in conducting needed operational analyses
- Missed opportunities to leverage funds by sharing data regionally
- Loss of management and public confidence due to inconsistent information issued by non-integrated systems.

## **1.3 Definition of Location Referencing**

Effective use of location referencing is needed to implement many of the tools transit agencies want to use to improve operational efficiencies and effectiveness. A technical definition of a location referencing method is the position of an entity relative to other entities or to some external frame of reference (e.g., latitude and longitude relative to the spherical geoid).

The term emerged from the data management field where an index is used to link spatial data within and between databases. The index or location reference is used to build relations between data concepts. The term also has roots in engineering surveying. Taken in this context, location referencing may be defined as stating the location of an unknown point by referencing it to a known point. For example, bus stop #356032 may be characterized or represented by numerous location referencing methods:

- 10 feet east on the south side from the corner on Elm intersecting with South Street.
- Near-side on the corner of Elm and South [on street, cross street]
- 25 Elm Street, Cambridge, MA 02140 USA
- route #10 mile 1.2, inbound direction

- <lat, long> WGS '84
- Northing, Easting MA State Plane
- link #58673901 on the XYZ digital (vector) base map
- Grid 45-245 on the ABC raster base map

## 1.4 Why a Location Referencing Best Practices Guidebook?

The industry is challenged by how to ensure that the location data itself is adequately collected, documented and maintained to meet evolving operational and data sharing requirements. A major issue is how to characterize location data throughout its life cycle and through its data flows so that spatial data can be shared and integrated. The industry needs to ensure that location-referencing systems meet the operational requirements of particular technology applications and of integrated systems. Transit agencies throughout the United States have been “learning the hard way” about location referencing problems and organizational issues pertaining to spatial data. They have also been creating some valuable solutions and best practices.

This Best Practices for Using Geographic Data in Transit: A Location Referencing Guidebook (LRG) clarifies the obstacles and provides best practices for transit on how to avoid the problems and facilitate successful spatial data sharing and use. It attempts to describe for managers and technicians the best practices for building and maintaining a robust spatial data infrastructure to support transit business needs including transit ITS applications. Many of the best practices described in this Guidebook also apply to other technology projects and data initiatives in Transit, not just to spatial data.

## 1.5 Guidebook Development Process

Transit staff, consultants and vendors who manage, use and share spatial information, developed the content of this Guidebook. They shared their expertise in a series of interviews, workshops and focused discussions on specific topics. Each draft of the Guidebook was reviewed in detail by transit agency staff and was open to comment by the entire industry. Appendix A describes the process for involving transit and spatial data experts in the development and review of this Guidebook.

Comments and suggestions for additions and revisions to the Guidebook should be addressed to:

Transit Standards Consortium (TSC)  
91 West Clinton Ave., Suite 2,  
Tenafly, NJ 07670

Or to the TSC Location Referencing Guidebook Committee Chair:

Bibiana Kamler  
GIS Manager  
Tri-Met  
4012 SE 17th Ave  
Portland, OR 97202  
Email: kamlerb@trimet.org  
503-962-7536

## **1.6 Goals and Objectives of the Location Referencing Guidebook**

The goal of the LRG Project is to assist the transit industry by facilitating more effective and unambiguous exchange of spatial data. The Location Referencing Guidebook is intended to be helpful, both directly and indirectly. Transit agencies, standards development efforts and vendors will benefit from the identification of business needs, issues and existing best practices. Given new tools and best practices information, transit agencies will benefit as spatial databases are built and expanded, as GIS and ITS applications are installed, as ITS applications are maintained and operated, and as spatially based analyses are requested.

The objectives of this Location Referencing Guidebook include:

- Identify barriers to transit spatial data transfer and management.
- Identify standards and “best practices” that minimize location referencing problems and support effective spatial data sharing—the Guidebook will provide the transit industry the most effective techniques and methods for representing, using, managing and exchanging transit features that possess spatial characteristics.
- Provide a high level “roadmap” for managers and additional detail for project managers and technical staff.
- Develop a comprehensive list of transit applications that use spatial data.

## **1.7 Scope of Location Referencing Guidebook**

The scope of the Guidebook is designed to address the needs of different audiences and to address problems and best practices from different perspectives.

### **1.7.1 *Different Audiences***

This Guidebook includes high level materials for transit managers, more detailed information for the Information Technology and GIS management teams and technical information for the technical staff. It is designed to help various types of transit staff, by providing both overview materials and detailed best practices. Specifically, it should help:

- Senior Management (identify best practices related to oversight and policy level issues; describe agency-wide view, barriers and responsibilities)
- Project Managers (identify best practices related to managing projects, systems engineering analyses, and implementation)

- GIS/IT/Data Management Staff (identify best practices related to technical development and maintenance)

Section 2 of the Guidebook, Road Map for Managers, is specifically written for senior transit managers. Section 3, Creating an Effective Spatial Data Infrastructure, is also worth a review by managers, although a few subsections get technical and may be skipped. Also the start of each Section contains a brief overview section that is written for all audiences, both managerial and technical.

### **1.7.2      *Transit Feature Data Emphasis***

The art of representing, reporting and managing a transit object is one of the most critical components of successful location referencing and transit spatial data sharing. One key to minimizing ambiguity during the transfer of spatially-based transit objects is ensuring the robust representation of transit features. Therefore, this Guidebook places a heavy emphasis on the creation and management of transit feature data.

Since transit features are created and collected by transit agencies, they may be subject to policies and procedures, or best practices, promoted by the organization. In defining the scope of what should be included in a Location Referencing Guidebook, transit experts in spatial data management identified areas where best practices may be described within the context of the data management and transit life cycles.

The Guidebook's best practices for transit features may assist transit with:

- Identifying transit program elements that use spatial data
- Understanding specific application quality ("fit for use") and referencing requirements, that is, developing the linkages among the various applications that use the data;
- Categorizing a transit feature as a primitive spatial feature
- Defining a common set of transit features, common names and formats used throughout the organization, including
  - General core spatial data (e.g., street network base map, census data boundaries)
  - Transit core data (e.g., routes, zones, facilities, ridership)
  - Non-core data (e.g., security incidents)
- Identifying specific persons with responsibility for managing the datasets;
- Identifying specific procedures for handling the spatial data as it flows through the life cycle (e.g., collection, translation to base map, etc.)
- Understanding data life cycle needs (how the data changes as it is collected)
- Identifying types of meta-attributes that capture data quality and procedures

### **1.7.3      *Different Perspectives***

While this is primarily a Guidebook on best practices for location referencing, effective use of location referencing requires both good transit systems engineering or project development practices and good spatial data management practices. Therefore this Guidebook covers best practices in those areas, because they are required in order to implement best practices in location referencing.

The complexities of spatial data sharing also suggests that the Guidebook should address different perspectives:

- Spatial data management (how to describe, build and associate the data),
- Standards (how to describe, share and communicate the data among the applications), and
- Institutional/policy (how to organize this within your property).

What this document is **not**:

- Primer on Geographic Information Systems or Database Management Systems
- Primer on surveying or maps
- Primer on data management
- User manual on a GIS application
- Procurement document for GIS or GIS services

## **SECTION 2    ROADMAP FOR MANAGERS**

### **2.1    Section Overview**

This Roadmap for Managers Section is intended to assist senior transit managers in enabling their staff and their transit organization to effectively acquire, manage and use spatial data to enhance transit operational effectiveness. Spatial data, such as the geographic location of bus stops, routes and the street network are used throughout transit, such as in customer information, planning, scheduling and dispatch. In particular, Intelligent Transportation Systems (ITS) applications, such as Automated Passenger Counting (APC), Automated Vehicle Location (AVL) and Itinerary Planning, require good quality spatial data to operate correctly. An extensive list of spatial data uses in transit and Geographic Information System (GIS) applications by transit business area is included in Appendix G to show the breadth of spatial data's value to transit.

The support and guidance of transit managers is particularly important with respect to spatial data because of the potentially high costs of mismanaged spatial data acquisition, maintenance and use. Transit is a very mobile business with many factors that move and change. The street network changes with growth, construction and detours. Transit routes, stops, facilities, employees and customers all frequently move locations. Data update and maintenance efforts can be very inefficient and costly in a transit agency if they are not well managed.

#### **2.1.1        *Purpose***

This section provides a high-level checklist to assist transit managers in effectively leading and supporting agency efforts to:

- Maximize the effective use of spatial data
- Minimize costly data/system redundancies and the resulting reconciliation efforts
- Foster data integration and sharing
- Establish spatial data as a key element of transit's data infrastructure and assets
- Minimize location referencing errors that cause operational failures and limit access to accurate information
- Be in compliance with the Federal Transit Administration's (FTA) National ITS Architecture Policy on Transit Projects

Most information management and technology projects will also benefit from following the "best practices" and roadmap elements presented in this section.

#### **2.1.2        *Problem***

Transit agencies are heavily dependent upon spatial data for operational effectiveness; knowing the location of transit vehicles, incidents, facilities and other resources is core to many transit business areas. Unfortunately, this means transit is particularly vulnerable to problems and issues with spatial data. These problems can cause capital project cost overruns, create costly operating inefficiencies and limit the effectiveness of the agency's information infrastructure.



Some of the potential problems are technical in nature and require technical solutions; however, many are at the policy or institutional level, where transit managers can make a difference.

Transit management leadership and support are needed for the success of new ITS and spatial data development projects. Effective decisions about spatial data usage are becoming more challenging with the prevalence of off-the-shelf GIS products, the embedding of GIS tools and maps in ITS applications, and the increasing demands for regional data integration with other agencies.

Transit management also needs to actively support effective information systems and data maintenance. The budget and resources that are expended to create a successful project implementation phase can be lost or wasted if the data are not appropriately updated and maintained. Poorly maintained spatial data can have an extremely negative impact on operational effectiveness. Similarly, investments in advanced technology applications such as ITS may not be fully realized if the technology is forced to operate using inaccurate or incomplete data. In some cases, incomplete or inaccurate data can create additional problems beyond not realizing the value from an investment.

Technology investments will inevitably be made, with or without management leadership and strategic guidance, but they will be more effective with the informed direction and support of transit management.

## **2.2 Role of Transit Management**

Transit managers rely on the successful operation of one or more automated systems; therefore they have a role in ensuring the existence of effective information systems and data. The transit General Manager and the head of Information Technology have particular responsibility for ensuring that an integrated, agency-wide approach is taken for developing data and information systems solutions. The development and maintenance of the spatial data and information management infrastructure in transit are just as important as the physical infrastructure. Active input from transit managers in other business areas is also needed to ensure that their data needs and business objectives influence the structure and content of data and systems. All managers have a role in supporting informed technology decision making and information integration, to ensure that the spatial data meets the business needs of their divisions.

Key roles of the transit management team are to ensure a common vision, communicate goals and priorities, be champions of integration, provide oversight and support staff. It is also the responsibility of the management team to ensure that the spatial data and related information systems explicitly support the operational needs of the agency; that technology does not assume a life and purpose of its own.

## **2.3 Roadmap for Managers**

The Roadmap for Managers consists of nine key steps. Implementing the steps will help ensure successful decision-making and operations by improving spatial data usage in analyses, maps,

customer information and ITS applications. Each step's importance and key actions to be taken by managers are explained below. A partial implementation of the key steps is unlikely to be effective. All of the steps need to be undertaken, although not necessarily in the order presented here.

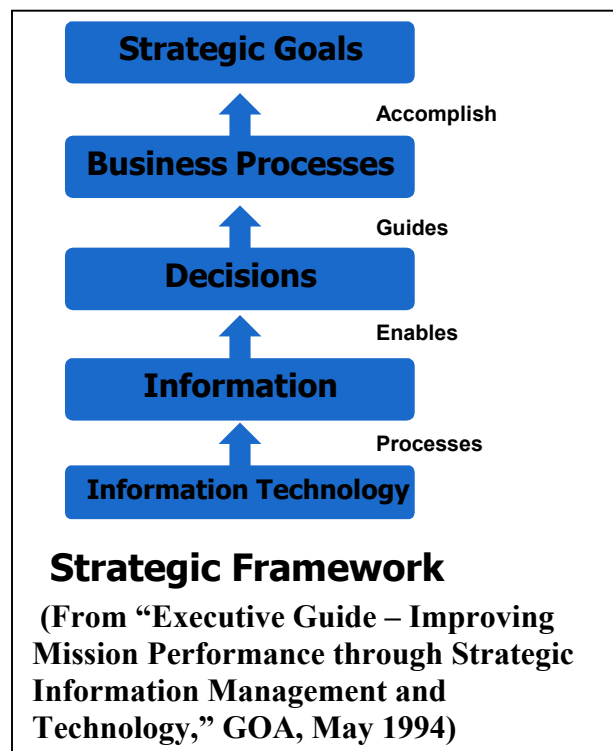
Some of these steps may seem familiar from the management and systems engineering literature. Effective enterprise-wide use of spatial data, and indeed of all information technology, requires a successful integration of traditional management practices with good systems analysis.

### **2.3.1      *Ensure Business Goals Drive Technology Investments***

The transit agency's general business goals should drive the agency's information and technology investments. Technology must support clearly defined business goals, such as "increase ridership" or "manage resources efficiently." Figure 2-1 shows a Strategic Framework that identifies the relationship of an agency's goals to its business processes, decision-making, information resources and information technology. It illustrates the following two points:

- An agency's mission and goals should drive technology investments.
- Information technology provides the support for effective decision making and achieving business goals.

**Figure 2-1 Strategic Framework**



Managers can help ensure that business goals drive technology and data investments by:

- ✓ **Approving technology projects that are tied to the agency's business plan.**

✓ **Developing and clearly communicating the agency's vision, mission and goals.**

In particular, project managers, IT and GIS technical staff should be briefed on the agency's goals and objectives. Over the course of a technology project, the technical staff makes countless decisions that can influence the extent to which the project meets the agency's objectives. Knowing the agency's goals and objectives, in addition to the project's technical requirements, can improve a project team's decisions.

### **2.3.2 Create an Enterprise-wide Approach**

Transit managers can save data, system and maintenance costs by promoting an enterprise-wide approach to data and systems. By looking at technology projects and datasets across an agency, redundant development and maintenance costs can be eliminated and systems interfaces can be simplified. For example, it is not effective to have staff who are responsible for service planning, customer service, AVL and APC all acquiring and maintaining their own separate base maps and route networks.

An industry "best practice" is to define data used by many groups as agency-wide "core data." Core spatial data, such as the street network, Bus Stop Inventory and routes, should be developed once and maintained without redundant updating. This is not to say that all core data must be developed and maintained in-house. There are strategic advantages to purchasing data, or to developing data in-house, but in either case, there should only be ONE authoritative set of core data for the agency. Data exchange among internal agency business units should be seamless. Managers can support an enterprise-wide approach by:

- ✓ **Consistently communicating a vision and expectation of integrated data and systems.**
- ✓ **Setting expectations for the definition, development and availability of "core data."**
- ✓ **Helping break down barriers between transit business areas and encouraging collaboration (create incentives and remove disincentives).**
- ✓ **Providing direction to staff to produce and maintain one "core" base map and one route network that can be distributed to other business applications and information systems for use.**

Note: Business areas may need to enhance or customize the data further for their application, but savings are still achieved by building on a common, core dataset.

If you have not previously instituted an enterprise-wide approach, it may be wise to do an inventory of interfaces, which may uncover some significant redundancies and opportunities for improvement. Similarly, you may need to do an inventory of spatial datasets within your agency if you have not already taken a "core data" approach.

### **2.3.3 Foster Consideration of Regional Partners**

Finding partners to share technology development costs makes good business sense for both operational and financial reasons. For example, scarce funding can be leveraged and riders can benefit from integrated services across transportation modes and service providers. A common street network, that is accepted regionally, facilitates data sharing among transportation providers and government agencies in a region. In addition, the recent new guidance on National ITS Architecture and Standards Conformity issued by the United States Department of Transportation (USDOT) FTA advocates shared systems and data development within a region. Valuable information about the policy can be found at the following web site:

<http://www.its.dot.gov/aconform/aconform.htm>

Managers generally have better regional contacts and connections than technical staff. As a manager, you know whom to contact in the region; your technical staff may not. Transit managers can assist technology staff and project managers by:

- ✓ **Helping to identify potential regional partners.**
- ✓ **Participating with the upper management team(s) of the other organization(s) to clarify the goals and expectations for the joint project and to make needed policy decisions.**
- ✓ **Helping establish effective communication and conflict resolution procedures.**
- ✓ **Setting a priority for your project staff to coordinate with regional partners and emphasizing that regional coordination must be considered as a part of new spatial data initiatives.**

For more information on regional coordination requirements, please see the FTA guidance. The guidance documents have detailed steps for developing regional data sharing plans and systems architectures. You may also wish to contact your Metropolitan Planning Organization to find out about existing regional data sharing initiatives. Your “region” may be a state – several states have started initiatives to create “base map” data for public agency use throughout the state.

### **2.3.4 Align Planning Efforts**

Transit has many planning obligations and requirements that directly or indirectly affect spatial data acquisition, availability, integration and costs. If planning isn’t coordinated at a strategic level and among different projects, many spatial data sharing problems can occur. Active management leadership is necessary to successfully leverage project funding and develop integrated data and systems. Without active management, department-level technology investments tend toward isolated islands of data and applications with redundant costs. Redundant costs are an immediate problem; redundant, incompatible spatial data can be a much bigger and longer-term problem.

The two most important ways managers can help align planning efforts and minimize problems with location referencing are by:

- ✓ **Verifying that ITS projects are consistent with the Regional ITS Architecture per the Federal Transit Administration's National ITS Architecture Policy on Transit Projects**
- ✓ **Requiring that ITS and IT strategic planning efforts are integrated.**

### **2.3.5      *Budget to Protect Investments***

New technology systems offer extraordinary opportunities for cost savings and operational efficiencies. They also bring new challenges with regard to systems integration and maintenance. The capital and operating budgets both provide opportunities to protect technology and data investments and to minimize ongoing operating costs. These benefits can be obtained by:

- ✓ **Setting expectations for coordinating projects internally and externally to leverage project budgets.**
- ✓ **Using cost/benefit and alternatives analyses to improve investment decisions.**
- ✓ **Ensuring that operating budgets are adjusted to handle the maintenance and update requirements of new projects or datasets such as the base map.**

### **2.3.6      *Increase the Effectiveness of Technology Procurements***

The procurement process is another opportunity to reinforce strategic initiatives and standards with respect to spatial data development, maintenance, integration and reporting. Procurement can also help the transit agency avoid redundant data, systems and processes. Management can increase the effectiveness of procurement efforts by:

- ✓ **Requiring Requests for Proposals (RFPs) and contracts to reflect consistent data policies and standards.**
- ✓ **Requiring intellectual property rights in contracts that allow easy, cost effective data sharing.**
- ✓ **Encouraging procurement teams that include key stakeholders (e.g., IT, the business client, ITS, legal and procurement).**
- ✓ **Requiring the specification of data integration and interface requirements.**
- ✓ **Evaluating and grading potential vendors/consultants on their ability to deliver data that are easy to maintain and integrate.**

### **2.3.7      *Require Education on Technology Options and Invest in Training***

For effective data sharing and minimizing costly system and data redundancies, knowledge is required of both current technology options and the transit agency's overall information needs.

A poor decision by inadequately informed or trained staff during a project development phase can often be more costly than training classes. Targeted, “short course” training can often be cost-effective. Transit managers can reduce risks and maximize opportunities by:

- ✓ **Requiring feasibility studies and alternative analyses to ensure that the transit agency understands the options available before a decision is made to proceed.**
- ✓ **Ensuring that systems and datasets are documented so that resources are known.**
- ✓ **Verifying that training on data management tools and techniques is budgeted.**
- ✓ **Arranging for information technology staff to learn more about the transit agency and its core information needs (e.g., via tours, speakers at staff meetings, videos).**
- ✓ **Funding training for system developers, maintenance staff and users.**

### **2.3.8      *Influence Strategic Technology Decisions and Implementing Standards***

Technology decisions have policy and operating impacts on a transit agency beyond the obvious impacts on capital and ongoing operating costs. Performance within transit is affected by the availability of technology elements such as an enterprise-wide GIS, widely accessible core data, adequate documentation and standards. These elements, and others, affect the ability to implement applications, the amount of time required to complete data analyses and maps, the type and quality of the analyses, and the skill levels of the staff needed to do the work.

General, agency-wide technology projects and initiatives serve as the foundation for many IT and ITS projects, but they frequently get little overall management attention and support. The general projects need advocates, critical assessment and strategic guidance just as much as projects specific to a business area within transit. All transit managers can improve the quality of strategic technology decisions by:

- ✓ **Voicing the agency’s general needs, in addition to project-specific needs, as they relate to technology.**
- ✓ **Supporting general technology improvements such as GIS and a standard approach (across projects) to report generation and map development.**
- ✓ **Allocating staff time for defining areas that need standards and for implementing the standards and procedures.**

[Note: These standards can be either internal standard operating procedures for using technology, such as using one GIS product throughout the agency, or standards that are formally recognized by organizations, such as American Public Transportation Association (APTA), National Transportation Communications for ITS Protocols (NTCIP) or the International Organization for Standardization (ISO).]

Again, managers play a vital role in focusing priorities and staff on the business requirements of the agency and on the need for information for decision making. Technology is a valuable tool that must be used appropriately to achieve the desired results.

### **2.3.9      *Provide Ongoing Oversight and Support***

Management support of technology projects throughout the project life cycle is critical. This includes the planning, budgeting, specification, assessment of alternatives, implementation and operations phases. Managers can provide ongoing oversight and support by:

- ✓ **Asking questions (e.g., Do you have an integration plan? Are the data documented? May I see the documentation?).**
- ✓ **Setting expectations for clear answers that can be understood by non-technical people.**
- ✓ **Reinforcing agency goals, including integration of systems and data.**
- ✓ **Helping to resolve competing priorities among different business functions when they affect a project.**
- ✓ **Supporting staff with recognition and encouragement as they work in a complex and changing profession.**

## **2.4      Think Beyond**

This Roadmap for Managers Section can also be useful for project managers and technical staff. Section 3, Creating an Effective Spatial Data Infrastructure, provides more detail on some of the key steps in the Roadmap. Managers can continue to improve their understanding of the spatial data sharing issues and how to avoid them, by reading portions of Section 3, and the Overview at the start of each Section of the Guidebook.

## **SECTION 3    CREATING AN EFFECTIVE SPATIAL INFORMATION INFRASTRUCTURE**

### **3.1    Section Overview**

Efficiencies in acquiring, maintaining, distributing and using spatial data can only occur when an effective spatial information infrastructure is in place in an organization. The infrastructure to support effective use of spatial data should include elements such as an enterprise-wide strategic planning framework for spatial data, readily accessible core data, effective database management access methods, and standards and procedures for maintaining data quality.

An enterprise-wide approach is particularly important because spatial data is used in so many different transit business areas, as shown in Appendix G. Further, looking across the transit industry, conflicting spatial data initiatives have been started in different organizational units such as Geographic Information Systems (GIS), Information Technology (IT), Intelligent Transportation System (ITS) and engineering. The different needs and initiatives should be coordinated from an agency-wide perspective.

This section elaborates on some of the elements of a spatial data infrastructure. It is primarily targeted toward management and technical staff in the GIS, IT and ITS areas. In some cases, this Section provides more detail on some of the items highlighted in the preceding Section 2: Road Map for Managers.

#### **3.1.1    *Problem***

In addition to developing the base map and overlaying transit features, a number of general information technology, information management and project management issues affect the successful sharing of spatial data and the ability of an organization to minimize location referencing problems. Some of the underlying obstacles to building an effective spatial information infrastructure stem from:

- Lack of common data definitions and a data dictionary
- Multiple and inconsistent datasets of core information (e.g., multiple bus stop inventories, different address formats and different base maps used by different applications or stored in different data sources)
- Inability to identify quality, accuracy and lineage of spatial characteristics of datasets
- Incomplete and imprecise definitions and descriptions of location referencing methods used by core data
- Confusing or difficult linkages to the spatial data and other datasets when doing reports

Many transit procurements are dependent on implementing standards and meeting FTA policies on conformity with the National ITS Architecture and Standards. Location referencing is a key area of concern within regional architecture planning and standards implementation because much of transportation's core data is associated with spatial characteristics and location referencing methods. Location referencing issues are not separate from the information



infrastructure; they are a key factor in enabling information integration and interoperability, and generating the information needed to further transit business goals.

### **3.1.2 Purpose**

A key objective of this section is to reinforce the need for an agency-wide approach to spatial data. A second objective is to describe the spatial data infrastructure elements needed within the Transit Information Enterprise, that is, "...the collection of core transit data and information that services the Enterprise, along with the needed applications, hardware, tools, policies, and procedures." [NTI course: Managing Information for Success: Your Role in a Changing Business, Glossary]

<b>Definition: Transit Information Enterprise</b>
---

The collection of core transit data and information that services the Enterprise, along with the needed applications, hardware, tools, policies and procedures.
---

Many of the details are described in subsequent sections:

- Establishing procurement practices that protect core transit data (Section 4)
- Standardizing on a single base map (Section 5)
- Defining a corporate transit feature dictionary and defining an enterprise reference data model (Section 6)
- Establishing and enforcing a data quality process and internal data standards (Section 7)
- Assigning responsibility for data (Section 7)
- Leveraging (and protecting) existing and regional application and data investments (Sections 9 and 10)
- Documenting the data life cycle through metadata (Section 11)

This section introduces spatial data infrastructure components such as a strategic framework, planning for growth, identifying and developing core data, developing a graphical user interface for spatial data, database access methods, implementing standards and improving project management and life cycle approaches.

## **3.2 Develop Infrastructure with a Shared Vision and Goals**

Transit agencies have a vision and goals to guide performance. The spatial data infrastructure should be developed to support the vision and goals of the agency.

### **3.2.1 An Integrated Business Enterprise**

Transit management has always been motivated to optimize performance within a business unit such as vehicle maintenance. Given increased operating pressures and more competition for limited financial resources, transit also needs to look for ways to optimize across business units. An agency-wide perspective is increasingly important for transit's success.

An integrated business needs integrated and shared data. This sharing of data is dependent on the organization being able to provide core and integrated data. It requires building an information infrastructure that considers:

- Links to business roles, goals and processes
- Accountability for data, data integration and data sharing
- Planning and budgeting for overlap and tradeoffs (understand opportunities)
- Incorporation of data management policies, procedures, standards and tools
- Systems engineering approach for procuring new systems/applications

In developing the information infrastructure, a transit agency may also consider:

- IT/ITS Capital Procurement/Strategic management planning
- Regional partners for sharing data and costs
- Use of tools related to National, Regional and internal (ITS) Architectures

### **3.2.2      *A Strategic Information Framework***

Successful business strategies and effective decision making depend on good information. The Transit Information Enterprise, along with effective information management techniques, is a critical part of the foundation for producing good information and business success.

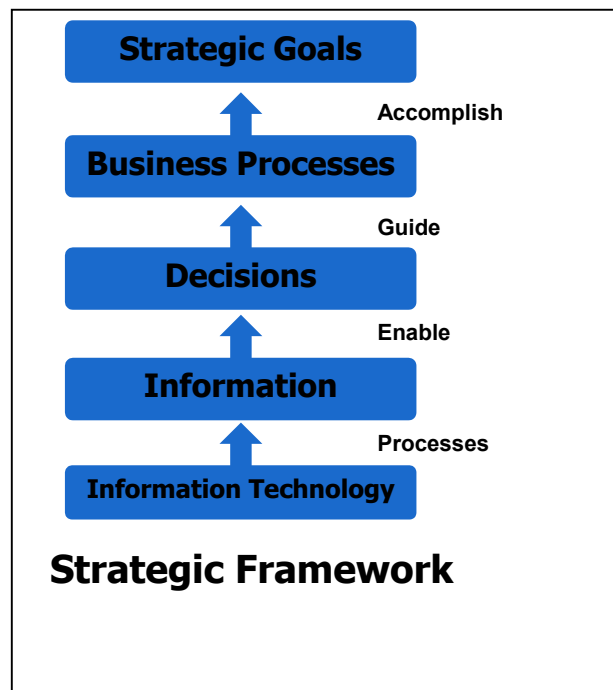
The Transit Information Enterprise should support an integrated business. The enterprise-wide approach defines “corporate” or core transit data that serves the whole agency. The base map and location of transit features are core datasets that serve broad agency needs. Performance measures and historic datasets must also be preserved with their spatial foundation. Preservation of spatial domain datasets that illustrate changing conditions (e.g., new streets, long-term detours, increased distance between time points, added stop points) are essential for planning and service implementation.

Data within the enterprise needs to be linked to support an integrated business. Prior to purchasing an application, the strategic goals an application meets should be well understood. The types of business decisions that are made in order to meet those business needs, and the information necessary to make those decisions, should drive the information technology that is procured. For example, if your executive management needs information on missed pull-outs, and you purchase an Automated Vehicle Location (AVL)/Computer Aided Dispatch (CAD), you will definitely want to identify and collect the information necessary to report “missed pull-outs.”

The “strategic framework” shown in Figure 3-1 was developed in the “Executive Guide—Improving Mission Performance through Strategic Information Management and Technology” [GOA, May 1994]. The strategic framework shows the linkages between agency goals and information technology. For example, a goal of improving service will generate a process to review on-time performance. Decisions that guide improving service are supported by answers to questions such as “How well are we doing today?” The information that enables answering that question is schedule adherence data, and the information technology that enables the collection of that data is AVL. Many of the business processes that deal with service quality

need information that is based on location. As a consequence, goals may be directly traced to the need to improve location referencing methods.

**Figure 3-1: Strategic Framework for Building an Enterprise Information System**



### 3.3 Plan for Growth

An information technology management organization should be guided by a well formulated IT strategic plan. An IT strategic plan describes the approach for managing transit's information technology and resources. The plan will consider the physical architecture, hardware, software and networking solutions for the organization. It may include a description of a target architecture composed of the application/data and physical architectures, and initiatives, processes, and maintenance activities related to sustaining and enhancing the current systems.

Spatial data and location referencing have been adversely affected when IT strategic plans do not include ITS or GIS. ITS and GIS should be included in the IT Strategic Plan. GIS is sometimes omitted because it may reside elsewhere, such as in planning. Sometimes ITS is not included in the plan because some applications lay outside of IT's responsibility, for example, the on-board bus devices or rail control systems. It is a mistake to leave ITS systems with on-board bus devices, such as AVL, annunciators and next stop signage, out of the IT planning process because they are information-driven and related to spatial data.

The ITS architecture considers an integrated transit architecture to include mobile vehicle subsystems, remote devices, customer services and management activities. All parts of the integrated transit architecture generate or consume information. As such, significant

dependencies may be ignored if ITS and spatial data issues are not considered during the strategic planning process.

The strategic plan impacts all applications and data sources, in particular, the base map and GIS applications and functions. For example, when describing transit business architecture functions, key functions should include “base map updates,” “scheduling data updates,” “bus stop inventory updates.” Changes to the base map will impact location references of transit feature data. Changes to transit feature data will impact the attributes (performance and configuration data) that are associated with them. The IT Strategic Plan should incorporate those change functions in describing initiatives and migration strategies to new applications and platforms.

As a best practice, include the following in the IT Strategic Plan:

- Develop management procedures for the base map and core data “authority” datasets
  - Identify core datasets
  - Develop management procedures over the dataset life cycle (e.g., acquisition/ generation, storage, processing, transformation, integration, visualization, update, archiving, etc.)
- Identify application dependencies based on data flows such as: “base map updates,” “bus stop updates,” etc. from core datasets
- Identify and leverage data integration requirements across multiple applications and procurements
- Develop procedures for data archiving; ITS data archiving should follow the ASTM Archived Data Management Systems (ADMS) Guidelines where. (ASTM International is a global forum for the development of consensus standards. See reference (Section 3.10) below.
- Integrate transit vehicle systems and remote customer device functions into the IT architecture (these assets also generate core information that must be integrated into the information enterprise).

### **3.4 Managing Spatial Relationships: Core Data and a Data Model**

#### **3.4.1 *Identify and Develop Core Data***

In transit, there are a number of commonly used data elements, or core data, that merit special consideration. These data elements are used in many automated systems and in many mapping and reporting efforts. Core spatial data examples include bus stops, routes, street names and the street network. These elements are used by many, but should ideally be maintained once.

Information driven applications require service, asset and other core information to operate effectively and accurately. The quality of the information must be consistent with applications that perform complementary functions. As such, there is a great need to ensure that data across an organization is the same, updated simultaneously and used consistently. Organizations also try to optimize deployment costs and maintenance resources. As such, rooting out redundancies, replacing manual with automated processes, and leveraging application interfaces supports the development of a complete and consistent spatial information infrastructure.

Spatial data, and ensuring interoperability of location referencing methods, is dependent on the information infrastructure developed by the Information Technology organization. The management of spatial data is a cooperative effort among end-users, GIS and IT staff.

Critical roles related to the development of a centralized dataset are the data “steward” and a Database Administrator (DBA). The data steward is typically a user who is responsible for the quality and validation of a particular portion of the dataset content. Typically, the DBA organizes and integrates the spatial data and provides automated functions that work on the overall spatial dataset. They are usually responsible for associating location reference methods with transit core data and ensuring that the base map is maintained. The data stewards and the DBA ensure that the relationships among a single dataset and other datasets are validated and consistent (also referred to as “referential integrity”).

As a best practice, the steps for developing a core data repository for spatial data are as follows:

1. Standardize around a single base map that is the source for the most current street network information (See Section 4 for a discussion on selecting a base map.)
2. Identify core data used by many applications and organizational units; gather requirements (attributes and representation needs) from users and stakeholders; (repeat process when additional applications are specified)
3. Identify key dataset owners or “stewards.” Their responsibilities include collecting/generating, updating and verifying the quality of their particular dataset. If possible, select data stewards from those sections of the organization that will be most motivated to ensure that the data are complete, correct and timely.
4. Determine data dependencies (integrated datasets, update frequencies, field survey needs, relationships)
5. Establish and implement a data model that supports spatial relationships and dependencies
  - a. Describe the spatial integrity procedures that should be implemented in the database
  - b. Transit spatial feature relationships tend to be complex (i.e., many-to-many relationships), and the model may evolve into “spaghetti” relationships
  - c. Not all data models support all applications; different views and interfaces may need to be developed to deliver model components in another format.
6. Develop linkages between the spatial core data and other enterprise-wide core data through a data model.
7. Build generic interfaces and tools based on the data model for data access, analysis and reporting.
8. Hire a trained Database Administrator. The databases cannot “run” on their own; they require trained personnel to manage planning, development and maintenance of the systems.

### **3.4.2      *Building a Data Model***

A robust data model that defines the relationships between transit features can support analysis, reporting and ITS needs. Due to differences in organizational structures, work policies, and cultures at various agencies, defining a single relational data model for all of transit is very challenging. A data model is often helpful for defining how applications will share information, and how data will be linked for analyses and reporting needs.

Developing a comprehensive data model that supports all of a transit agency's data is a daunting task. Few organizations have been successful when they try to develop and implement a comprehensive model. Limited data models that represent parts of the enterprise linked by key entities is a "modular" approach that has been successful. This approach assumes an organization has a good idea about the key relationships among its data. Issues related to pieces of the model are analyzed and developed in an evolutionary approach. Using a reference model such as one developed by another transit agency, Geospatial One-Stop (GOS) (limited to geospatial data), or one developed by the Europeans (i.e., TRANSMODEL) will help mitigate the risk of this approach. (Section 3.10, Resources, includes references to GOS and TRANSMODEL.) Other ways to limit the impact of changes to your data model are discussed in Sections 6.4.3 and 9.4.

One of the major advantages of developing a data model is the implementation of internal referential integrity procedures. These procedures perform format checking, verification of identifier uniqueness, business rule checking (e.g., "patterns must terminate with valid time points"), and updating/flagging of related features (e.g., if a bus stop is retired, the pattern that contains an ordered sequence with that bus stop is identified as changed, too). The integrity checking may apply to both attribute and spatial relationships of the transit feature. These internal consistency rules can help ensure that the location referencing attributes are accurate and translate appropriately between applications.

## **3.5      Develop a Common Graphical User Interface for Spatial Data**

A transit agency's investment in spatial data has diminished value if the data can not be easily located and used. A standard, agency-wide graphical user interface (GUI) can significantly improve the ability of systems analysts and users to locate available data, determine its appropriateness (see metadata) and run reports or create maps. By standardizing the look and feel of the query and report/map generation process, less training time is required and results can be produced faster. A standard, agency-wide GUI can also facilitate communication between end-users.

The common graphical user interface for spatial data should be very similar in form and approach to the GUI for non-spatial data in the agency. Ideally data queries should easily access both spatial and non-spatial data. Some of the principles for an effective GUI include the use of standards, simplicity, accessibility, web access, simple forms for routine data queries, and a "layered" approach from simple to more complex to accommodate a range of users.

### **3.6 Implement Standards**

Standards improve quality and reliability, while reducing training time for staff, costs and data integration complexity. Standards can be ones developed by formally recognized standards organizations—National Transportation Communications for ITS Protocols (NTCIP) or International Standards Organization (ISO)—or internally defined standards (e.g., using internally-defined Extensible Markup Language (XML) tags or standard operating procedures (e.g., data update procedures). Benefits can be achieved by implementing a wide variety of standards, such as for:

- Data definitions
- Location referencing approaches
- Base map
- Address standards for filling out incident reports and other forms
- Metadata
- Data quality procedures

A list of recommended standards related to spatial data are included in the Resources , Section 3.10.

In addition, adoption of other standards will support the development, evolution and strategic approach of the transit information enterprise and its spatial data component. These standards may cover areas such as:

- Platform configuration requirements
- Database management system
- Software languages
- Computing architecture approach (e.g., J2EE, .NET)
- File and graphic formats
- Server configuration and portioning policies

### **3.7 Use Effective Database Access Methods**

The computing architecture in which the spatial infrastructure is embedded will support or hinder the ability to share data. A centralized set of core data may be implemented in a single database or distributed databases. Database access may be provided through generic procedures or specialized Structured Query Language (SQL) statements. The implementation of a transit data model is itself unique to each agency; even the transit reference model developed by the TRANSMODEL working groups (see reference below) is implemented differently by different public transport agencies. The unique implementations pose barriers to agencies sharing data, sharing applications and leveraging resources. The design work and decision making can be complicated. A trained Database Administrator is recommended to do the work.

Typically, when a new application is introduced to an enterprise information system environment, the Database Administrator, users and vendors discuss adding new fields to a specific table, and changing or revising relationships required by the new application. Changes to the database organization are reflected in the data model; changes to the data model impact the integrity of data interfaces that use SQL statements to access data in the database. A change to

the data model will invalidate a SQL statement that relies on that relationship to access information. Many times, the relationships are embedded in application programs that access the information. There are techniques to “hide” the logical relationships of the data model between data from applications and mitigate the impact of changes to the database. These techniques are discussed in Section 9.5 on The Impact of Client/Server Models.

### **3.8 Maintain Spatial Data Quality**

Spatial data must first be developed or acquired that meets the data quality needs of the users. After the data are available, the quality must be maintained. In addition, users and systems analysts must understand the quality of the data, as discussed in Sections 7 (on Quality) and 11 (on Metadata).

When developing or acquiring spatial data, it is not sufficient to simply specify the accuracy, completeness and other elements of quality that are required by the users. Tests of the data quality must be completed; some tests designed to hunt for specific types of errors and some random tests. In Section 5, the subsection titled “Guidelines for Evaluating Commercial Base Maps” provides some specific ways to test the quality of base maps.

To maintain the spatial data quality over time, well planned data maintenance and quality-testing procedures need to be developed and implemented. Again, random quality checks on the data should be performed and diagnostics routines and reports should be developed and monitored. A key component of a successful data maintenance effort is the involvement of many people in the quality effort. Agency staff who see or hear about spatial location errors or mapping errors (e.g., incorrectly located bus stops can be noted by facilities maintenance staff, operators, customer assistance staff who handle complaints, or staff members that use the transit system) should know where and how to submit information about the problem. Section 7 of this Guidebook is devoted to spatial data quality issues.

### **3.9 Improve Project Management and Life Cycle Approach**

Although not technically part of the spatial data infrastructure, project management approaches can impact the development of an enterprise-wide spatial data infrastructure, spatial data sharing and location referencing issues. Ideally, projects should maintain both a business unit perspective and an agency-wide perspective. Sometimes, it is for the good of the agency to expand the scope of a project and incur schedule delays in order to have the project serve more than one client and leverage the agency’s technology investment. For example, multiple Global Positioning System (GPS) antennas might be purchased and installed on a bus, if coordination did not occur between projects such as AVL, Smartcard, Automated Passenger Counters and Annunciators. In addition, better products result when all the project life cycle phases are completed, particularly the alternatives analysis phase.



### 3.10 Resources

#### ITS Standards

For more information on ITS Standards, see the USDOT Standards Program web site:

<http://www.its-standards.net>

#### Metadata Standards

Federal Geographic Data Committee. (1998). Content Standard for Digital Geospatial Metadata. FGDC-STD-001-1998.

[http://www.fgdc.gov/standards/documents/standards/metadata/v2\\_0698.pdf](http://www.fgdc.gov/standards/documents/standards/metadata/v2_0698.pdf)

ASTM E2259-03 Standard Guide for Archiving and Retrieving ITS-Generated Data. Advisory includes an overview of several projects that demonstrate how agencies are using the standard to develop and implement Archived Data Management Systems (ADMS). The Advisory can be found at:

[http://www.its-standards.net/Documents/ADUS\\_Advisory.pdf](http://www.its-standards.net/Documents/ADUS_Advisory.pdf) (PDF)

[http://www.its-standards.net/Documents/ADUS\\_advisory.htm](http://www.its-standards.net/Documents/ADUS_advisory.htm) (HTML)

#### Data Definition Standards

##### TCIP Data Elements and Messages

NTCIP 1400 Transit Standards Information Profile (TCIP) Framework, version 1.05

NTCIP 1401 Standard on Common Public Transportation Objects, version 1.03

NTCIP 1402 Standard on Incident Management Objects, version 1.03

NTCIP 1403 Standard on Passenger Information Objects, version 1.03

NTCIP 1404 Standard on Scheduling/Runcutting Objects, version 1.03

NTCIP 1405 Standard on Spatial Representation Objects, version 1.03

NTCIP 1406 Standard on On-Board Objects, version 1.03

NTCIP 1407 Standard on Control Center Objects, version 1.03

NTCIP 1408 Standard on Fare Collection Objects, version 1.03

#### Spatial Data Standard: Geospatial One Stop

The Geospatial One-Stop is a Federal government program to develop “building blocks” for a National Geographic Information Network. It is one of 24 e-government initiatives sponsored by the Federal Office of Management and Budget to enhance government efficiency and improve citizen services. Geospatial One-Stop will make it easier, faster and less expensive for all levels of government and the public to access geospatial information.

**“Geospatial One-Stop will provide tools for all levels of government—local, State and Federal—to work together more effectively to:**

- Facilitate sharing of geospatial information
- Improve planning for future investments in geospatial data
- Expand collaborative partnerships that help leverage investments and reduce duplication
- Collaborate on the development and implementation of standards that encourage sharing and use of best practices and advance implementation of the National Spatial Data Infrastructure.”

<http://www.geo-one-stop.gov/about/index.html>

### Standard Transit Reference Models

Transmodel is conceptual data model that describes the functions and relationship among public transport entities. The model was developed by the European public transportation community as part of the European standards development organization, CEN.

<http://www.transmodel.org>

## **SECTION 4    GENERAL PROCUREMENT GUIDELINES**

### **4.1    Overview**

A transit agency's procurement policies and practices can play a significant role in facilitating data sharing and minimizing location referencing problems. Enforcement of internal, regional and federal policies for procuring transit applications and systems promotes best industry practices. "Best practice" policies such as specification of standards, though potentially appearing as higher "up-front" costs, are promoted because they support planning for contingencies, long-range maintenance savings, and modularity of design.

The value of standards for enhancing flexibility and long-term functionality of equipment can be illustrated by an example pertaining to home entertainment systems. A television set is significantly less valuable and flexible without standard attachments for cable, DVD, VCR and stereo speakers. The physical connectors are based on standards and the digital signals support agreed upon protocols. The form and function are driven by market expectations.

#### **4.1.1        *Problem***

Procurement policies and procedures may help or hinder advancing best practices for location referencing. Most often, it is the lack of integration requirements and standards that cause problems. Missed opportunities can occur when procurement staff take a project-specific perspective, without balancing it with an additional agency-wide perspective, with respect to spatial data sharing requirements and standards.

Procurement efforts can be further challenged because transit applications must meet many demands and work in complex environments. Their system specifications can be difficult to describe and meet. The transit applications support long life cycles, they must share information with existing and legacy systems, and they must work within the organization's computing environment. With respect to spatial data, they must use the agency's base map, conform to its feature data definitions/formats, meet the agency's business goals for producing quality information, and leverage the agency's current and planned investments.

#### **4.1.2        *Purpose***

This section highlights for management and procurement staff, some of the key provisions and policies that should be included in procurement efforts for projects that use spatial data or generate data that are visualized on a base map. This section also briefly reviews the procurement requirements of FTA's National ITS Architecture Policy on Transit Projects. Many ITS projects rely on a base map, use transit data that possess spatial attributes, and share these data with other regional partners.

Additional topics reviewed in this section include:

- Intellectual property and data sharing issues
- Procurement provisions for data and interfaces

- Enforcement of industry and Information Technology (IT) Standards and internal policies
- Evaluating proposals

## **4.2 FTA's National ITS Architecture Policy on Transit Projects**

A relatively recent FTA policy affects transit procurements of ITS applications and spatial data. On April 8, 2001, the Federal Transit Administration National ITS Architecture Policy on Transit Projects went into effect. This policy requires that Intelligent Transportation Systems (ITS) projects funded by the Highway Trust Fund and the Mass Transit Account must develop a regional ITS Architecture, use a systems engineering approach during development and use ITS standards adopted by the United States Department of Transportation (USDOT). Key goals of the policy are to facilitate data sharing, improve integration between travel modes and service providers, and to leverage funding.

Procurements of affected transit ITS applications will require more attention to standards and the specification of data sharing relationships. The policy stipulates that:

“The final design of all ITS projects funded with highway trust funds shall accommodate the interface requirements and information exchanges as specified in the regional ITS architecture. If the final design of the ITS project is inconsistent with the regional ITS architecture, then the regional ITS architecture shall be updated.”

A key aspect of developing the Regional ITS Architecture is collaboration among internal and external organizations, which is discussed in more detail in Section 4.3.

In addition, to be in compliance with the policy, all eligible ITS projects require the execution of a systems engineering analysis that includes the following:

- A description of the scope of the ITS project
- An operational concept that identifies the roles and responsibilities of participating agencies and stakeholders in the operation and implementation of the ITS project;
- Functional requirements of the ITS project;
- Interface requirements and information exchanges between the ITS project and other planned and existing systems and subsystems; and
- Identification of applicable ITS standards

Many of these elements affect the content categories and level of detail in an ITS or spatial data procurement document. Section 8 of this Guidebook provides a more detailed look at the systems engineering requirements relative to spatial data.

## **4.3 Collaboration Among Internal and External Organizations**

Federal Transit Administration's (FTA) National ITS Architecture Policy on Transit Projects reinforces the need for collaboration on procurement efforts among key stakeholders such as IT, the business client, ITS, legal, procurement staff and regional partners. Roles and requirements pertaining to the development and/or maintenance of spatial data need to be defined.

**Collaboration among external organizations.** Transit staff should develop strategies for sharing base map resources with external organizations. Furthermore, they should establish formal agreements on the communications, security, data management and information exchange requirements for all interfaces, and the resources needed to implement those interfaces.

In particular, special consideration should be taken with respect to sharing transit features that include spatial attributes (location information). For example, in developing a Regional 511 system, the regional entity may desire to use a base map that encompasses a region larger than Transit's service area. For that reason, the region may be driven to purchase or develop a base map that is inconsistent with the transit agency's base map, and consequently with the transit information (e.g., bus stops, route alignments, etc.) that is associated with the base map. This decision will increase your cost for managing and maintaining key datasets. These hidden costs adversely impact the effectiveness of the regional initiative. Section 10 of this Guidebook discusses strategies for interagency coordination, including shared data development and management activities.

**Collaboration among internal organizational units.** Procurement can play a role in helping ensure collaboration between key stakeholders on the system requirements. Procurement policies can provide a final check to ensure that enterprise-wide system integration requirements are considered. IT projects and software applications seldom live in a vacuum in today's environment. Transit organizations must integrate information to improve decision-making and better meet business goals.

Best practices, that also help ensure conformance with the FTA policy, include:

- Identification of key internal and external stakeholders and their requirements
- Sign-off on the specifications, roles and responsibilities by impacted organizational units and stakeholders
- Procedures for multiple stakeholders to use common datasets, and if necessary, justification and a cohesive strategy for using duplicative datasets or base maps
- Estimated update procedures and resources needed for the base map and other datasets that are necessary to ensure proper operation of the application or system
- Inclusion in the specification of explicit descriptions of application interfaces and data sources that support or receive data from the system under procurement
- Implementation of formal project development procedures

For more information on Federal grant procurement policies, see Chapter 23 of the Grants Management Workbook, which is available on the Internet at:

<http://www.fta.dot.gov/office/program/gmw/23ITS.html>

## **4.4 Intellectual Property and Data Sharing Issues**

Some aspects of intellectual property (IP) are often overlooked in the procurement document, resulting in painful, costly consequences for transit. Intellectual property rights with respect to updating, accessing, using and sharing data need to be addressed. If the IP needs of the agency are not specified, transit can be stuck with redundant inputting of new routes and stops,

redundant updating of base maps, delays in receiving updated datasets, and an inability to share data with other applications, business units, transportation agencies, and the public.

Spatial data (either base map or transit information with spatial attributes) may be the goal of a procurement effort or it may be hidden in an ITS system procurement specification. For example, some turnkey ridematching, trip planning or automated vehicle location (AVL) systems are sold with embedded base maps. Agencies have run into difficulties when they focused their procurement requirements on system functionality and had too few requirements for the data.

There are different categories of spatial data that may be procured:

- Base map
- Electronic sets of transit feature data, including field survey data
  - Bus stop and route inventory (including benches, signs, other amenities)
  - Track and alignment
- Transit application generated data
- Related base map datasets (e.g., demographics, landmarks)

Transit may desire to share these datasets across the organization, with regional partners or on the Internet. Without appropriate attention to IP language during the procurement process, the licensing requirements may not allow agencies to replicate, extract or recompile data for other applications and analysis. Organization of the data storage (e.g., the back-end database data model) may be proprietary; consequently, presenting an obstacle to building interfaces between applications that must share information.

Contract imposed data restrictions can create other types of problems as well. For example, transit staff may have permission to extract data from an application and the expertise to enhance it, such as by adding new streets, transit facilities and landmarks. However, in some cases, if the enhanced data or “entities” are put back into the database, warranty problems may be triggered. Numerous transit and paratransit organizations have spent significant time upgrading street network and address data in their ITS applications, only to have most of their work lost when a new update to the file is provided by the vendor. Negotiating better working partnerships with the vendors during the procurement process can ease some of the inefficiencies in updating data files.

In the past few years, more attention is being paid to IP issues related to the application code because of the short life of some transit vendors. Transit agencies are realizing that they may own and operate software long after a vendor has gone out of business or has been acquired. As a result, they need to protect their business continuity by protecting their ability to update the software code and documentation over the long term.

Attention also needs to be paid to data IP issues in the procurement specifications. Data IP issues can impact flexibility of data use and distribution, efficiency of the data update process within the organization, ITS project implementation schedules and agency data integration needs. IP language in the procurement specifications should address the data generated and the data interfaces to and from the products.

Intellectual property law is still a developing field and intellectual property issues will probably continue to grow. If a transit agency does not have in-house legal counsel, outside legal expertise may be needed initially to help understand IP issues and develop contract language. If a transit agency has in-house legal counsel, they may be interested in training seminars on intellectual property pertaining to data ownership and sharing. Technology implementation projects often support the training of team members on new software, so investing in the skills of in-house legal staff can also be appropriate and valuable in the long-term.

#### **4.5 Procurement Provisions for Data and Interfaces**

A key recommendation for the procurement of any transit application that uses spatial data (base map or transit feature) is to include provisions in the procurement contract that clearly define the data input and output requirements and the responsibility for developing any needed interfaces. Identify the data dependencies in advance and identify other products to leverage in order to produce additional benefits. Responsibilities for the vendor, organizational unit and project team need to be defined.

The interfacing requirements are based on business needs. The requirements may address how often service is changed, how often bus stops are changed, what the impacts of the changes are on the customer and what data enhancements stakeholders may need. Application procurements should assess three sets of interfaces: input data, output data and maintenance cycles for input data. Technical considerations related to these datasets are included in Section 9.

Procurement may ensure that these provisions are incorporated in the specifications by the following:

- Identification of the data both needed and provided by stakeholders
- Identification of the format that data should take (e.g., a data dictionary)
- List of responsibilities for developing interfaces for each of the data input and output (including base map, application generated transit information, and input transit information)
- List of responsible parties for (input and output) data maintenance and management

#### **4.6 Standards and Internal Data Policies**

The costs of managing duplicative datasets, inconsistent systems, and incompatible hardware can drain an agency's valuable resources. Development and use of internal computing policies and standards such as coding languages, PC configurations, network connectivity, and data descriptions serve to mitigate inefficiencies. The FTA recognizes the need to promote standards. In 2001, FTA enacted the National ITS Architecture Policy on Transit Projects to help encourage the transit industry to adopt standards and facilitate data sharing and regional transportation solutions.

The enforcement of the following standards and internal data policies will facilitate spatial data sharing, support productivity, decrease cost overruns and improve on-time completion of ITS applications:

- Standardize and use a single base map as a source to support all your applications
  - Specify that applications read a standard format for updated base map files
  - Standardize on a set of location referencing methods and association tables (e.g., street alias table)
- Define standards for inputting addresses into applications (e.g., location of security incidents)
  - Develop street addresses for transit centers and park-and-ride lots
- Use a single set of surveyed locations for transit data such as bus stops and patterns
- Use an agency-wide data dictionary of transit data descriptions

#### **4.6.1      *The Open-Source Movement in Government***

Government agencies are taking a closer look at open-source software (OSS). Products that claim open-source must adopt a licensing agreement based on the principles of open source (see [www.opensource.org](http://www.opensource.org)). With open-source software, the source and compiled code are offered to the user, the user may change or add functionality as long as the product offer uses the same licensing agreement as the original code. Open-source software also allows free redistribution, which is very important for a public organization that must share information with other agencies and the public. Obviously, open-source software offers many advantages to government agencies that have restricted tax dollars and broad requirements for openly distributing public information. Examples of open source software include Linux, MySQL (database), Apache Web Server, Preprocessor Hypertext Protocol (PHP), and many others. Distribution of applications based on OSS does not require costly or complex licensing agreements with commercial vendors; it can be distributed freely.

The movement is gaining influence because many governments in the US and abroad are adopting language that encourages government agencies to review OSS products. Some countries like South Africa and Australia have passed legislation to adopt OSS. Although, states like California and Oregon have been unsuccessful in adopting resolutions to formalize preference for OSS, organizations like NASA and jurisdictional governments within States are adopting policies to analyze technology trade-offs based on standards-based, open technology architecture, as well as open source software licenses. Some organizations include selection criteria that weights standards-based, open architecture, and open-source software licenses higher than proprietary systems. Tri-Met is one transit organization that is currently developing language to include in all information technology systems engineering studies and procurements that sets out a justification for prioritizing systems based on standards, open architecture and OSS. Through the FTA Policy on ITS, the Federal government is currently advocating inclusion of language prioritizing system selection based on standards and standards-based architectures.

Note: The OSS movement should not be confused with the free-ware software efforts that produced GNU products (e.g., compilers). An article discussing the differences between the approaches may be found at <http://www.gnu.org/philosophy/free-software-for-freedom.html>



## 4.7 Evaluating Proposals

The evaluation criteria for rating proposals should reflect the transit agency's need for spatial data that are easy to maintain and integrate. The evaluation elements should consider, among other items, the proposer's ability to:

- Support the transit agency's spatial data standards and location referencing needs.
- Be easily compatible with the agency's database and reporting tools, as well as other relevant internal standards.
- Allow the transit agency to import or update spatial data as needed.
- Allow the transit agency to create or extract spatial data as needed.
- Permit transit to share data with different applications, work groups, regional partners and customers.
- Support interoperability and a standards based approach to computing.
- Protect the transit agency's ability to access and update the spatial data and source code in the event of a vendor being unable or unwilling to continue supporting the transit agency.

## 4.8 Resources

### **Annotated Bibliography of Technical Standards**

Best Practices in Procurement Manual (published by FTA):

<http://www.fta.dot.gov/library/admin/BPPM/ch8.doc>

### **Other Resources**

"The Road to Successful ITS Software Acquisition." Federal Highway Administration, USDOT, July 1998.

### **Open Source Software:**

<http://www.opensource.org>

Link to Oregon's Open Source Software Bill:

[http://pub.das.state.or.us/LEG\\_BILLS/PDFs/HB2892](http://pub.das.state.or.us/LEG_BILLS/PDFs/HB2892)

## **SECTION 5 THE BASE MAP**

### **5.1 Overview**

A “base map,” required by many transit business functions and Intelligent Transportation System (ITS) applications, is a Geographic Information System (GIS) database containing features such as streets represented as centerlines with address information associated with them. This Guidebook uses the term “base map” to mean the foundational spatial and attribute data to which the transit data are referenced. Section 5.2, Definitions, discusses this in more detail. Section 6 discusses transit features and how to reference them to the base map.

#### **5.1.1 *Need for a Base Map***

A base map is essential for every transit agency that wants to understand where its riders, potential customers, facilities and vehicles are located. An effective transit agency base map provides consistent transit data across applications within the agency and allows efficient sharing of transit data with other transit and non-transit agencies. Further, a primary agency base map eliminates costly and time-consuming redundancies in data development, maintenance, testing and reconciliation. Location referencing problems are minimized if best practices are employed when acquiring and maintaining a base map.

Transit agencies use a base map to store and manage information about its service area. The base map is also required to generate a spatial representation of bus routes, which historically, have been maintained as tables in scheduling software in the form of a pattern. Transit agencies cannot perform many service and planning analyses, nor run most ITS applications, without its service footprint being related to a base map. That is, the base map is the foundation for managing transit features and understanding the spatial relationships among the transit features, such as routes and stops. The key issues associated with “relating” transit objects to the base map that will be discussed in Section 6. Base map requirements, which are covered in this section, are key to ensuring that the base map is fit for use across the transit enterprise.

#### **5.1.2 *Industry Pressures and Agency Inefficiencies***

The evolution of transit application software for planning, scheduling and ITS, in concert with emerging transit business requirements, have put tremendous pressure on the development and maintenance of a base map. In the struggle to quickly implement new planning tools and ITS systems, transit agencies often inadvertently fall into the trap of paying multiple times for the acquisition of redundant base maps.

As a secondary consequence, the transit agency also ends up paying multiple times for the management and maintenance of multiple versions of the base map. When new streets are added to the service area or changes occur, time and money must be spent placing the changes on the various base maps. Transit routes and facilities, when they change, must also be linked to the different base maps. Redundant, inconsistent base maps frequently produce conflicting data and results, which can create time-consuming data reconciliation efforts and public relations

problems for the transit agency. More on incremental updating of a base map; versus scheduled quarterly or annual updates is included in Section 5.4 and Section 9.

One factor that contributes to agency inefficiencies with respect to base maps is the lack of a strategic framework or architecture for the agency's spatial data infrastructure. Another factor is transit business units independently implementing ITS applications such as automated vehicle location (AVL), ridematch or trip planning without realizing that they were each acquiring a base map with their software.

### **5.1.3      *Problems in Managing a Transit Base Map***

The key problems that arise related to managing a transit base map include:

- No consensus exists in the transit industry about what should be included in the base map because individual transit applications may have different requirements. For example, a Trip Planning application may require pedestrian access to streets, whereas, a scheduling application may not.
- Changing business requirements lead to new data requirements.
- New application technologies (e.g., Global Positioning System (GPS) and AVL) can require increased spatial accuracy.
- Regional mandates for data sharing require integration mechanisms across multiple agencies.
- Commercial transit applications may use proprietary data formats for spatial data. These proprietary formats create expensive and cumbersome updating and translation processes. For example, some applications may have different rules for character type or length, which may make data preparation difficult and costly.

### **5.1.4      *Purpose***

Transit agencies can benefit significantly from guidelines that improve base map acquisition and maintenance practices. Base map investments can be leveraged and agency inefficiencies can be reduced. The purpose of this section is to provide transit agencies with strategies to streamline the acquisition and maintenance processes by minimizing the number of base maps and number of base map versions.

This section discusses the issues related to acquiring and maintaining a transit base map. Best practices are related to the following:

- Establishing a single base map as core to the agency spatial data infrastructure
- Defining the essential data elements for a base map that can serve transit
- Evaluating the decision to buy or build a base map for transit
- Evaluating commercial base maps and base map vendors
- Evaluating the decision to participate in regional data sharing

Since the base map is essential to using and sharing transit spatial data, many of the other Guidebook sections will elaborate on other issues and best practices that impact the base map.

## 5.2 Definitions

The term “base map” is used in GIS application areas to mean different things depending on the application domain. Even within the transit application domain, there is confusion and lack of consensus among transit practitioners about what is meant by a “base map.” Part of the confusion lies in the fact that commercial applications include different feature types in a single data storage called the “base map.” More confusion arises when application vendors describe the base map differently. For example, an application vendor may classify all linear features (streams and streets) in a single base map layer.

This Guidebook uses the term “**base map**” to mean the foundational spatial and attribute data. In the context of transit applications, a base map is, at a minimum, a GIS database containing streets represented as topologically accurate centerlines with address information associated with them. In addition, depending on the transit agency’s requirements, the base map can include other physical, real-world transportation “objects” of interest for transit operations and analysis such as rails, ferry lanes, and possibly pedestrian features. These features are represented in a GIS as points, lines, or polygons with associated attribute information. The base map can include travel and turn impedances. A base map used for transit applications may also satisfy the requirements for other transportation applications. Transit objects are associated with the base map’s spatial and attribute information.

The term “base map” was selected because of its common usage in the transit industry. In this Guidebook, the base map is the dataset that can be used to produce many map illustrations, depending upon the point of view, which displays the street network and landmarks. After transit features are associated with the base map, transit objects such as bus routes, bus stops, and ridership may also be illustrated relative to the base map. The base map supports both cartographic output (maps) and analysis (scheduling, route planning, etc.).

Some transit practitioners use the terms “transportation network” or “centerline file,” rather than the term “base map.” For some, “transportation network” is more descriptive for their agency because it suggests a navigable database that can be traversed, whether by bus, car, bike, rail or foot. However, a network representation only requires connectivity not necessarily geometry (shape). The term “centerline file” is another synonym but there is no consensus about what should be represented by each centerline and some centerline files do not include network connectivity. “Base geometry” and “base cartography” are other terms that are sometimes used to describe aspects of the base map.

Many other layers of spatial information can be referenced to the base map, such as digital orthophotos, zip codes and census boundaries. Aerial photos provide an easy-to-interpret backdrop for trip planning applications. Orthophotos and satellite imagery provide additional data sources to evaluate and validate an agency’s base map and to support the placement of attributes.

### 5.3 Related Transit and ITS Applications

The base map should be thought of as a core, enterprise-wide resource. As a centralized spatial data repository, the base map provides consistent transit data and reference framework for virtually all transit applications. The spatial data can be managed and maintained efficiently in the agency base map, then transformed and exported as needed to other systems. The base map is core to all aspects of the transit business. The numerous transit information systems that rely on data from the transit base map include:

- Automated Vehicle Location
- Automated Passenger Counters
- Trip Planning available over the phone or via the web
- Real-time bus location on the web or at stop locations
- Ridematching
- Paratransit for demand-response reservations and dispatching
- Mileage tracking used for reporting purposes and vehicle maintenance
- Safety and security systems for incident tracking
- And many other applications that are listed by transit business function in Appendix G.

### 5.4 Recommended Practices

#### 5.4.1 *Single Base Map*

One of the key recommendations of this Guidebook is that transit agencies acquire and maintain one base map that is used to support the needs of applications across the enterprise. All additions and changes to the street network would be done once on the base map. As stated in Section 5.3 above, the base map should be thought of as a core, enterprise-wide resource. As a centralized spatial data repository, the base map provides a consistent reference framework for virtually all transit applications and required transit data. The spatial data may be managed and maintained efficiently in the agency base map, then transformed and exported as needed to other systems.

Although referenced in this section, transit information is not an integral part of the base map. Transit information may be thought of as an “event” or “attribute” of the base map. Section 6 deals explicitly with transit information as geographic “features” and discusses issues related to associating them to the base map.

For the agency’s base map to be successful, the appropriate stakeholders and their requirements need to be defined, understood and met (See Section 5.4.2). Staff that manage and maintain the base map need to understand that different business functions and applications may not be able to use the base map without transformations. Although there may appear to be multiple maps, GIS applications and base map representations in the agency, ***they all derive from a single source***, stored in a centralized location. This single base map is used for recording changes and additions, and when updated, the replication process is repeated to support a consistent representation to end-users and applications throughout the organization. This replication process requires implementation of a synchronized data maintenance schedule whenever the base map is updated. Maintenance synchronization is discussed in Section 9.

### **5.4.2      *Define Stakeholders and Their Requirements***

Before assessing base map options, the base map stakeholders and their requirements need to be defined. This involves looking a bit into the future for agency requirements. The task is extremely important because the cost of replacing or upgrading the base map later to meet unforeseen requirements may be more costly than the initial acquisition. Not only will the agency pay for the new or upgraded base map but also for transferring all locationally-referenced objects over to it.

Often, funding from a variety of in-house or regional ITS projects can be pooled and leveraged to create a better base map or lower individual costs. Similarly, other regional partners such as local governments and utilities may bring their requirements and possible funding and staffing contributions.

The most important step in defining the requirements for a base map is defining functional requirements for the underlying business operations. First, define what you want to DO with the spatial information before specifying technical elements such as accuracy and attribution. For example, a transit agency may want to locate a bus stop on the correct side of a dual-centerline road. The spatial accuracy of the bus stop placement would have to be at least one half of one lane width, to ensure that the stop is placed on the correct side of the road. The transit agency may also require that the bus stop be linked with the correct side of the dual-line, not just identified by spatial coordinates. If an agency wants to use a base map to support more efficient Paratransit dispatching, the agency will need to know the actual address ranges for the blocks, not just theoretical address ranges. If GPS-based AVL or APC systems will be used with fixed route transit, the level of accuracy will likely need to be greater in areas where the buses travel than in other areas of the base map.

### **5.4.3      *Essential Data Elements for a Transit Base Map***

There are no formal standards or consensus about what constitutes a base map for transit. Every transit agency has some unique operating requirements and some unique features in its operating environment. However, best practices have been identified in transit, regarding the development and maintenance of an agency base map.

Elements of the base map form the basis for a wide range of GIS applications, the support of ITS applications and asset management and analyses through location and temporal referencing (e.g., time points, linear location, geocoding). Transit agencies should understand each of the following essential elements of a transit base map.

- A transit base map is required to generate a spatial representation of bus routes, which historically, have been maintained as tables in scheduling software in the form of a pattern. In addition to producing maps for a wide variety of audiences, many transit agencies now need to run ITS applications that require spatial data. Using the spatial representation of the routes and stops represented on a base map, in conjunction with temporal data such as schedule and time point information, a linear reference, such as

distance from the intersection or from the last stop point, can be generated. The linear reference is needed to support ITS systems such as Trip Planning, which need to produce itinerary data for customers.

- A navigable database of the street network is needed to serve as the base transportation network. To facilitate maintenance and data sharing, each transportation mode (e.g., roadways, bike paths, ferry ways, railways) may be managed in a separate data layer. The separate layers may be linked together through connectivity at common “junctions” or “nodes.” All transfer points should be nodes but not all nodes are transfer points.
- Each transportation layer must have a topological (link/node) representation with travel (direction of traffic flow). Turn impedances (topological connectivity) which describe allowable turns are also highly desirable. Topology (connectivity) may be an attribute of the geometry. In that case, rules for the centerline representation must be consistent with the topologic representations that are needed for transit applications. Some agencies use one centerline for two directions of travel on undivided roads and two centerlines, one for each direction, on divided roads.
- The base map must support location references to elements such as zip codes, municipalities, parks and address ranges.
- The base map must support spatial referencing methods, including linear referencing (distance offset), coordinate referencing in an established XYZ mathematical representation or map projection, and map-matching (positioning GPS coordinates in latitude/longitude onto the transportation network).
- The base map will not remain functional over time without procedures and rules for collecting and maintaining the transit base map data.

Examples of transit features that can be related to the base map include route patterns, bus stops, time points, transit centers, park-and-ride locations, points of interest and fare zones. Points of interest are the common origin and destination points for transit passengers. These include landmarks, parks, hospitals, shopping centers, businesses, monuments, and schools.

#### **5.4.4      *The Decision to Build or Buy a Transit Base Map***

Whether a transit agency is going to buy or build a base map, it needs to consider whether it is proceeding with or without partners for cost sharing and consistency. First, it is not recommended that a transit agency build a base map on its own. Building a complete and accurate base map is time consuming and costly. Today there are multiple map vendors that sell base maps. In addition, other possible public partners exist that need a base map similar to transit, such as local governments, metropolitan planning organizations, utilities and E911 service providers.

Since there are regional data sharing requirements and benefits, it is recommended that transit partner with other local jurisdictions and transportation agencies to acquire and use a common base map where possible. The common base map can either be purchased from a vendor or developed locally.

Transit should negotiate to protect its ability to share the data as needed to meet all of its business requirements. It should also carefully negotiate its ability to update the base map or acquire updates in a timely, flexible and efficient manner. The choice of a purchased base map versus a locally built base map that is in the public domain, depends on the capabilities of transit's potential local partners, the value offered by the vendors and transit's agency-wide business requirements and resources.

### **Regional Data Sharing**

Regional data sharing efforts are agency consortiums with formal partnerships among the agencies to leverage data resources and facilitate data sharing. By way of partnering, regional transportation data can be shared at one or both of two levels:

1. Data sharing across multiple transit agencies within a region.
2. Data sharing across governmental agencies within the region, including transit. The resulting regional base map serves transit and transportation applications as well as other applications employed by local governments and utilities in the region.

The benefits to transit of regional data sharing are huge. In some cases, regional data sharing is required as a part of the Regional ITS Architecture. The benefits and issues presented in Section 10, titled Intra-Agency and Interagency Coordination, must be considered before proceeding with a plan to implement or upgrade a base map. Section 10 discusses motivations for regional data sharing, benefits, potential data partners, impacts on transit, some case studies, and critical success factors. In addition, relevant interagency data sharing benefits and issues are discussed.

### **Historic Precedence for Building Base Maps**

Prior to the commercial availability of navigable databases, in-house or contracted development of a base map was the only options. These base maps were built from public data sources such as U.S. Census Bureau TIGER line files and digital orthophotos.

Decisions regarding the life cycle of a base map—acquire, maintain, retire and replace—are driven by changing transit business requirements and the demands of new application technologies. To a much lesser extent, expansion and evolution of the street network or the necessity to share data across government sectors affect the decisions. According to Mike Berman at King County Metro, “We replaced our circa 1980s base map in 1993 with an enhanced U.S. Census Bureau TIGER line file. We are about to replace that one with an in-house photogrammetrically-developed network. The primary driving factors for the most recent replacement are the need to increase absolute accuracy for GPS-based AVL, and completeness beyond the fixed-route footprint to support Trip Planning and ADA adherence.”



The real world is constantly changing as a result of construction, changing names, changing topology, expansion of high occupancy vehicle (HOV) lanes (important for FTA reporting), new or changed addresses, or even road closures as a result of natural events like earthquakes or floods. Transit agencies must have a long-term maintenance program in place with dedicated staff and resources to gather and encode network changes. This is especially true if transit agencies rely on publicly sponsored data because the agency must assume the maintenance responsibilities. Areas with rapid growth can be especially burdensome because of large maintenance costs.

Some of the benefits of using a locally-maintained transit base map that is in the public domain are the following:

- The data maintainer is closer to the source and can incorporate business processes to ensure accurate and timely maintenance (i.e., road closures, vacated streets, construction).
- Responsibility for data is in the hands of the maintainer who is also the main data user.
- Maintainers will incorporate data features specific to transit because it is in their best interest as a data user, provider and maintainer.
- Local control over the frequency of updates—allows real-time updates if so desired.
- Local control can avoid loss of edits between major updates which sometimes occurs with vendor-provided updates.
- Local ownership of the data helps ensure that the data can be shared and used freely.

### **Motivations for Purchasing a Commercial Base Map**

Today when transit agencies want to acquire a new base map, they have a choice. Agencies may invest in the redevelopment of a base map or purchase a navigable database from a map vendor, such as NavTech, TeleAtlas, GDT and other local providers. Several motivations lead to the latter choice.

- The agency can compress the base map development schedule by starting with quality line work of the streets and the associated network connectivity provided by map vendors.
- The agency may not have the staff expertise to build the geometric, geographic, and topologic representation of the roadway network. Simply put, some agencies say, “We are in the transit business not map-making business.” By purchasing a map, the agency can reduce, but not eliminate, the need for this type of expertise on their staff.
- Once the base map is acquired, an agency has several options for ensuring that the original investment is protected through an ongoing update and maintenance process. One option is for transit to assume the long-term responsibility for developing and maintaining its street network in-house. However, most agencies that use a commercial base map do so to reduce the burden of long-term maintenance. Depending on the vendor and desired data format, currently most transit agencies obtain semi-annual or quarterly updated copies of their base map including line work and attributes. If a transit agency better defines its data update needs, they can work in partnership with the vendor to develop alternative data update relationships. Finally transit can form partnerships with other organizations or the vendor to share the update burden and make the update process better meet requirements while being more cost and time efficient.

Experienced transit agencies report that commercial map vendors tend to provide better coverage for urban areas or public roads, than for rural areas or “private” roads. There may be a qualitative difference in positional accuracy and navigational capability in rural areas (less accurate and less capable) and there may be less frequent maintenance due to inherently less growth in rural areas. Similarly, coverage of the navigable roadway network through areas such as shopping malls, airports, and university campuses may not be as complete or accurate. The transit agency may need to generate and attach missing geometry and attributes in those areas.

A contract agreement to receive updates may not fully relieve the transit agency from the burden of maintaining the base map. The responsibility of the transit agency partner is to report errors and omissions as guidance for the map vendor to improve quality of the database. For that reason, these agreements are referred to as “partnership.” Alternatively, some agencies may choose to do no maintenance in areas of slow change or where applications do not require greater currency or accuracy. However, the no-maintenance strategy is not ideal in most cases.

The major impact to Transit is that when an updated version of the base map arrives. Agency personnel must re-evaluate the base map at each update cycle to determine if changes were captured or whether geographic features such as transit features, landmarks and other datasets (“patches”) may need to be related to the base map again due to improvements to the foundational centerline and attribute data.

#### **Budget for Replacing the Transit Base Map**

Transit agencies have to anticipate and budget for base map replacement because the process can take a long time to accomplish and can be very expensive. Experienced agencies warn that getting a new base network, whether purchased or developed, is only part of the full task. Agencies will need a significant budget for enhancements to the new network to support the business needs. Mike Berman at King Count Metro says, “There will be missing geometry and incorrect or missing attribution. If routes reference the base map, and some of them go through shopping center parking lots, large business parks, or university campuses, then the base map will probably need to be modified to accommodate these. And for trip planning, customers want door-to-door trip plans necessitating that the accuracy and completeness go beyond the fixed route footprint.”

#### **5.4.5      *Lessons Learned by Transit Agencies in Using Commercial Base Maps***

When transit agencies purchase commercial base maps they frequently end up with a number of operational problems that might have been avoided.

Transit agencies need to clearly understand and document their data update and maintenance requirements, taking into consideration their needs under temporary and emergency conditions. These requirements need to be clearly communicated to the vendors. Transit also needs to work with the vendors to develop contracts that serve and protect their needs. In the absence of stating and negotiating maintenance requirements with vendors, transit gives up flexibility, efficiency and control over the maintenance and update of the transit base map.

Some of the problems encountered by transit agencies are listed below.

- Sometimes, the agencies cannot update temporary and permanent changes on the fly.
- Often, agencies cannot make changes to support “what-if” analyses for planning in the future.
- Changes and fixes that are completed by transit agencies are frequently not persistent. The improvements get lost in the next update that is provided to transit. A best practice is to not change the vendor’s database. Instead, use a layer overlay of agency data to handle persistent errors. Experienced agencies have developed efficient data procedures to deal with persistent anomalies and egregious errors, and to add missing geometry to the vendor’s data.
- Semi-annual or quarterly updates may not be frequent enough for some transit applications such as paratransit and trip planning, or in areas where the street network is constantly changing due to rapid growth.
- “Fixing persistent errors can be like swimming upstream.” Some persistent errors never get fixed because the burden of identifying, isolating, and communicating the extent of the problem is put on the transit agency reporting the error to the vendor.
- Transit agencies may not be able to use data updates in transit applications because the vendor data model is not compatible with the data models used in the transit application. Sometimes the problem stems from inadequate contract language with either the base map vendor and/or the transit application vendor.

Commercial base maps often do not provide complete coverage for the entire service area.

- In rural areas, commercial base maps tend to have lesser geometric and geographic quality, more missing line work, and more missing attributes than in urban areas. (The Utah Transit Authority discovered that commercial base maps for their region did not have sufficient network coverage and attribute information in rural areas.)
- The street network and geometry in a commercial base map may not show streets through shopping malls and industrial parks that are used by transit vehicles, which are needed for transit base maps. A best practice is to have transit agencies develop patches digitized from orthophotos and the associated data processes needed to update the base map.

Finally, deploying a new update is not a plug and play process. There is no such thing as a clean update to the transit base map. Agencies regularly experience problems with swapping in new transit data street geometry. According to Shirley Hsaio, Orange County Transportation Authority (OCTA) requires a transaction file to be delivered with each new update of the base map. The transaction file clearly documents all changes such as street length or ID. OCTA uses the transaction file to identify/evaluate changes that impact the agency’s transit data.

#### **5.4.6 Guidelines for Evaluating Commercial Base Maps**

Transit agencies recommend evaluating the quality of alternative base map databases before making a selection. The evaluation checks suggested in this section focus on roadway network geometry, topology and attributes as fundamental properties to support routing, AVL and GPS applications. (The base map may contain additional features, such as political boundaries and hydrography.)

- Select 25 subareas that require transit or paratransit service provision for evaluation. When selecting the subareas, include rural and urban areas, areas with recent construction as well as areas with no change in years, areas with complex interchanges and multiple lanes and areas with private roads.
- Investigate spatial accuracy and completeness. For each area, compare the base map to ground-truth (Most agencies use reliable surrogates for ground-truth such as digital orthophotos.). In older subdivisions, are there any missing streets? Do newer subdivisions and rural areas show up on the base map? How well do centerlines in the base map follow alignments in the orthophotos? For recent construction, does the base map reflect new alignments?
- Investigate attribute accuracy and completeness. Verify street names, addresses and other attributes provided by the vendor. Does the vendor provide travel direction?
- Investigate navigational topology (connectivity) to support routing. Verify that roads are connected and directed (one-way, two-way) correctly. Verify that connectivity (ramp configurations) at interchanges is correct. Verify that viaducts, tunnels, bridges are shown correctly. For bridges and viaducts, verify that upper and lower roads are not shown as connected. Upper and lower roads representing bridges and viaducts do not necessarily need to be topologically disconnected if turn restrictions are being maintained.
- Pre-determine the process for referencing agency-specific attributes (routes, time points, etc.) to the new base map. Compare these processes for each vendor's data model.
- Be prepared to make trade-off decisions between absolute (positional) and relative (connectivity) accuracy. For most transit applications, connectivity is more important. For AVL and other applications that depend on correctly relating geodetic (GPS) coordinates to the base map, positional accuracy is also important.
- Develop a method for handling single versus multiple centerlines. Some agencies use a single centerline for both travel directions on undivided streets while others use dual centerlines (or carriageways). Some agencies use dual centerlines when the divided highway has a median of at least 25 feet. Problems will occur at borders when integrating data between agencies or jurisdictions when one agency uses dual and the other single lines.
- Determine if the file has other transportation network information such as bikeways and rail, or attributes such as obstacles to walking and walking paths. These attributes may be required for customer trip itineraries and access analysis studies.
- Determine whether the data can be exported in the formats that are of use to the various transit applications and regional data users.

International standard organizations have been working for years to establish standard metrics for vendors to express map database quality. These efforts include ISO 19113 (Geographic Information - Quality principles), ISO 19114 (Geographic Information - Quality evaluation and procedures) and ISO GDF 4.0 (Geographic Data Files). Section 7 presents more details of these standards along with other considerations and practices for establishing and maintaining data quality.

#### **5.4.7      *Guidelines for Evaluating Commercial Base Map Providers***

Transit agencies recommend that the following questions and issues be considered when evaluating alternative base map providers.

- Consider the nature and business practices of each provider.
- How well does the vendor's existing base map meet your agency's business requirements? Does the vendor routinely provide data to applications with similar requirements? For example, a transportation map vendor is more likely to have a map that meets your needs, rather than a vendor who primarily sells demographic/marketing maps to sales research organizations.
- Is the vendor committed to developing an accurate (relative and/or absolute) network representation in your area? Does the vendor have other customers in your area?
- What are the vendor's data sources and processes for maintaining the base map? Does the vendor employ field staff (road data collectors) in your transit area? Is the geographic extent of the field staff such that major urban areas may be given higher priority?
- How frequently will updates be provided? Is the frequency interval sufficient? Is the frequency of updates compatible with business needs?
- If an urgent update is required, how fast can the vendor respond? What will it cost? Each transit agency needs to consider the definition of "urgent," both under normal operating circumstances and during natural disasters or attacks.
- Is the vendor data model compatible with the data model requirements of the agency's transit applications?
- How will the agency update transit application datasets when updated vendor data are received? Will there be additional vendor charges for any required data conversions? Can you negotiate a one-time price for conversion software rather than paying for each conversion as a service?
- Are link and node IDs persistent from update to update? (Changes in the link or node IDs may impact externally stored attributes that are associated with those IDs.) How are changes in the base map reported with each new update?

- What is the vendor's process for identifying necessary updates to the base map? How does the vendor know that a street name has changed? How does the vendor know that a street has been closed?
- What are the vendor's criteria for determining whether to encode temporary changes to the base map? For example, one vendor (Navtech) uses the following rule: If the temporary change, such as construction, has a life of six months or greater, then the base map is updated to reflect the temporary change.
- How will your transit agency report errors to the vendor? How will the vendor respond to and address the errors? Several vendors accept online requests to have the network updated in a particular area. A vendor may direct field staff to collect data to update the area with poor coverage. Another vendor may also accept change requests to correct errors. The vendor may attempt to verify the error by driving the roads (road data collectors) or using orthophotos before making the changes. What will be the lag time to see reported errors in updates?
- What is the license agreement? Will the transit agency be able to share data with other transit applications in the agency? What data will transit be able to share with other government agencies and the public? In what form?
- Will the transit agency be able to make timely changes to the base map in response to emergency situations (natural disaster or terrorist attack)?

## 5.5 Technical Discussion

### 5.5.1 ***Lesson Learned: Need for Unique Segment IDs and Links to Historical Identifiers***

This section lists key lessons learned on issues related to segment identifiers (IDs) and base maps. In some cases, vendors and/or transit agencies have applications and procedures that depend on segment IDs. For example, base map segments may be aggregated to form routes. The suggestions below are relevant in those cases.

However, some transit agencies, such as Tri-Met, are moving away from depending on base map segment IDs. Tri-Met uses spatial conflation tools to build some transit objects and stores x,y coordinates. Section 6.4 includes a discussion on how to build transit features and discusses some of Tri-Met's approaches.

- Before purchasing a map database, transit agencies should learn the vendor's business rules.
- Link and node IDs should be persistent from update to update; that is they should not change IDs from one update cycle to the next. Some transit agencies report that, in some

vendor's base maps, the segment (link) and node IDs for the roadway network change from update to update. If link and node IDs are not persistent, then transit agencies will have to match old with new link and node IDs so that agency-specific data can be propagated to the new base map. The matching process can be very time-consuming.

- When there is any change to a node or link, the old IDs should be retired (not reused). New or replaced links and nodes should be given new IDs.
  - Furthermore, transit agencies suggest that in addition to using persistent unique IDs, vendors should provide a report identifying historical relationship changes between versions of the base map. This report may be used to build historic queries using different base maps. The report should indicate which links and nodes were retired, replaced or added. This report would enable transit agencies to propagate agency-specific updates to the base map quickly and with less error. Presently without this information, agencies have developed spatial and non-spatial data tools to search and discover differences between base map versions. For each discovery, a human must perform a comparative investigation to understand the nature of the change. [Authors' note: We are not aware of a vendor who provides this type of reporting.]

### **5.5.2      *Patches***

If a transit agency receives its base map from another organization, new updated versions of the base map may only be available once a month, once a quarter or once a year. Between major updates, transit may have to “patch” the base map to fix errors in the street network or to add significant detours, such as when a bridge is damaged. If possible, a mechanism should be negotiated with the base map provider so that all the interim updates or patches done by transit are not lost. A mechanism should be developed that logs all the changes and when they were implemented. For interim patches, such as in the case of detours due to major construction projects, a system should be developed to alert transit staff when the patch should be removed.

A patch is a section of a base map that contains missing or corrected geographic and topologic information. The patch “lays on top of” the incorrect section of the map. It is registered to the map and all the edges and street network are matched. For example, if a base map does not include a private road through an industrial park, additional line work may be drawn within the industrial park. The patch will be overlaid within the industrial park area (using registration marks or “ticks”), and the base map street network will be connected to the street network drawn for the industrial park. A topological junction, such as a node, will be inserted to ensure topological connection with the existing base map.

### **5.5.3      *Hidden Costs of Proprietary Data Formats***

Some commercial transit applications (such as some scheduling and trip planning products) require the base network to be represented in a proprietary format. When agencies purchase these transit applications and want to use a regional or enterprise-wide base map, they frequently must purchase customized code that converts the agency's base map to the proprietary format necessary to run the applications. If a transit agency or map vendor changes the format of the base map, then new customized code will need to be purchased from the application supplier.

Costs for base map conversion can be significant in terms of dollars and time delay in deploying a new base map. In addition to paying for a base map, many transit agencies have to pay transit vendors roughly between \$10,000 and \$40,000 for the conversion routines. Unfortunately, some transit agencies pay again and again, because their contract with the transit software vendor is priced for each conversion of an updated base map into the proprietary formats. The \$12,000 conversion cost spent by one agency did not cover the cost to convert the base map data into the formats of any other applications, nor to run the conversion program again during the two-year update period provided by the base map vendor. Furthermore, the data format conversion process took two months. For UTA, the cost to translate semi-annual updates of the base map into the proprietary format for one transit application far outweighed the costs of the original map product.

When cost (dollars and time) to translate between data formats becomes prohibitive, then transit agencies are forced to conduct business while working with out-of-date, incomplete information and different update versions of their base map.

If a transit agency pays for a conversion program to input the most current base map into an ITS application, the transit agency should specify in its contract with the application vendor that the transit agency can run the conversion program itself, whenever a map update is needed. No additional, external costs should be incurred when an update is incorporated.

## **5.6 Think Beyond**

The evolution of base map data models and maintenance strategies, and their relationship to data sharing partnerships, and transit software applications should be guided by two currently unmet requirements:

1. The ability to implement real-time updates to commercial base maps. Satisfying this requirement will enable transit agencies to quickly adjust service in response to incidents such as traffic accidents, road closures, natural disasters, and homeland attack.
2. The need for implemented standards to facilitate the ability to translate between data formats used by commercial base map vendors and the formats (currently proprietary) used by transit application software. Satisfying this requirement will enable transit agencies to overcome the current labor intensive bottleneck for deploying updated versions of the base map to the various transit business applications within a transit agency.



## SECTION 6 POPULATING AND ORGANIZING THE BASE MAP WITH TRANSIT FEATURES

### 6.1 Overview

Once a base map is acquired, the transit features that compose the transit network must be associated with that base map. A *transit feature* is defined as “a real-world transit object or event (an event is an “action which occurs at a time, location or both) that is represented in a Geographic Information System (GIS) as a point, line or polygon with associated attribute information.” Transit features such as bus stops, patterns, vehicles and fare zones may be located by reference to objects in the real world. An *attribute* is defined as a characteristic of a transit feature. For example, attributes of a bus stop may include a bench, signage, lighting, timetable information and accessibility characteristics.

When the relationship between transit features and the base map is defined, a transit feature may become an attribute of a spatial feature in the base map. For example, a bus stop may be represented as an attribute of a particular link in the base map. The attribute is associated with the base map through a *location reference* or *location referencing method*. Then again, a base map feature or an attribute may become an attribute of the transit feature. For example, a bus stop may include qualities such as the width of the sidewalk, intersection street names, nearby landmark, and zip code.

This section discusses issues related to defining and managing transit features and their attributes, so that transit spatial data can be linked to the base map and effectively used. It assumes a relational database approach, as that is the foundation of most commercial GIS programs. Information on field surveys for collecting transit feature data and its attributes is not included in this section.

#### 6.1.1 Problem

Many transit agencies encounter significant problems with using their transit data because of the way they associated the transit data to the base map. The base map serves as the foundation for organizing a transit network. As a result, problems can propagate rapidly if data definitions and relationships are not set up using “best practices.” Many agencies encounter difficulties in the following:

- Positioning transit features on the base map
- Defining, building, organizing and storing transit features
- Establishing relationships between key transit features
- Associating attributes (static, dynamic and performance) to transit features
- Identifying the difference between a transit feature, attribute and event

In many ways, most commercial GIS packages function like a relational database, where each feature is an attribute of a node, link (also called vertex or edge, respectively), or series of links. As such, the transit feature must reference a node or link in order to be associated with the transportation network.

The method of relating transit features to the base map can impact the effectiveness of the spatial analyses and ITS applications. Integration with other transit features and transportation network attributes (e.g., curb cuts, permissible turns, and divided carriageways) are also affected by the way that the transit feature is related to the base map.

The two most common mistakes performed in “layering” transit features onto the base map are:

- Tracing a route over the base map line work
- Placing point features such as bus stops on the map using only latitude and longitude.

The transit features, in order to inherit the characteristics of the underlying base map, must be “attached” to the base map by directly referencing the base map network or the attributes associated with the base map. A trace is an overlay on top of the network and does not share the characteristics of the network (e.g., topology, road names, address ranges, etc.). The topology is necessary for many spatial analyses and ITS applications. Ridership and running time data cannot be flexibly aggregated and mapped without identifying the relationship of bus stops and time points to the routes. Similarly, the method of relating the transit routes to the base map can affect a transit agency’s ability to do corridor analyses. This topic is discussed in more detail in Section 6.4.1.

Placing a point in space (using only latitude and longitude) may be useful for some applications, but cannot support enterprise-wide data and analyses requirements. For example, when bus stops are identified by global positioning system (GPS) coordinates, but are not attached to the road segments, their location relative to the street network is ambiguous. This is an example where the location referencing method (latitude/longitude) is not appropriate by itself for the application. The location of the bus stop must be transformed or translated into a different location reference to support effective analysis. Operationally, you want to know where the bus stop is in relation to the street network and to various routes, rather than its location in relation to a mathematical model of the earth’s surface.

Transit features are used in many business applications and automated systems within a transit agency. Appendix G includes an extensive list of transit functions and systems that use transit features within a spatial context. One of transit’s challenges is that transit applications have differing feature needs in terms of data quality, attributes, relationships and location reference to a base map or the real world.

The location referencing method used in one transit application may share transit features with another application that uses a different location referencing method. For example, GPS-based on-board systems (e.g., automated vehicle location, automated passenger counting) require that a “trip” feature be defined as a series of events characterized by latitude and longitude and time. In contrast, trip itinerary planning applications require the trip feature to be defined within the context of a topologically-complete base map. As part of the “best practice,” a transit agency should implement a location table (discussed in Section 6.4.5) or another method of automatically transforming among the location referencing methods used within the agency.

### **6.1.2 Purpose**

The purpose of this section is to describe standards and best practices related to defining transit features and associating them to the base map. This section also examines issues related to managing and integrating transit features. This information is intended to help transit agencies avoid or minimize problems with spatial data. Even if your agency chooses to hire a vendor to associate features with the base map, it is important to know the common problems that occur, and some best practices, so that you can ensure that the delivered product does not have these common problems.

Section 6 discusses key transit feature definitions and attributes, after a review of location referencing definitions. The remainder of the section focuses on recommended practices for transit features and an introduction to an emerging standard that supports transit feature sharing, the Geospatial One-Stop's Transit Geospatial Reference Model. Issues pertaining to building a data model are identified. Some of the best practices include an approach for building a Location Table to support feature portability to different frames of reference and an approach for storing transit feature data.

Recognizing the variability in the industry in describing transit features, this document adopts the Geospatial One-Stop (GOS) definitions that are based on the Transit Communication Interface Profile (TCIP) definitions. The TCIP definitions were developed over several years by a wide range of transit professionals.

## **6.2 Location Referencing Methods Defined**

Transit applications use and create transit feature data and attributes, yet each application may require different location referencing methods. This subsection describes various types of location referencing methods that are used to associate a transit feature to a base map.

Among the types of referencing systems, the most common are:

- **Geographic:** a means of identifying a location by reference to a geographic or spherical frame of reference that represents the "real world." For example, latitude and longitude coordinates, state plane coordinates, or map grid references.
- **Linear:** a means of identifying a location by reference to a segment of a linear geographic feature (such as a roadway) and distance from some point along that segment. [ISO 15046] For example, mileposts, or offsets from a street intersection.
- **Attribute:** a means of identifying a location by reference to an attribute on the transportation network. Typical attributes include street names, addresses, and landmarks.

Transit agencies and vendors use different *geographic referencing systems* to define the location of features. The most common geographic referencing systems are:

- Latitude and Longitude (e.g., WGS '84 datum)
- State Plane Coordinates

- Modified Grid Systems (to reduce bandwidth requirements for radio communications), where the grid may be based on map pages, transit service areas assigned to road supervisors, radio tower coverage or vehicle depot/garage region

Transit agencies and vendors also use different *linear referencing systems* to define their linear networks. A significant characteristic of a linear network is that it includes topology. Linear referencing methods are characterized by an offset from a known location on the network. The type of network will drive the offset description. The most common linear referencing methods are:

- Offset from street intersection (on a street centerline file)
- Offset from route or pattern (on a transit network)
- Relative to intersection (near side, far side, mid-block near side, mid-block far side) (on a transit network combined with a street centerline file)
- Percent traversed from start of path (e.g., along a pattern, route, roadway, trip itinerary)

Some transit ITS applications depend on other base map *attributes for referencing locations*. Paratransit and itinerary trip planning applications commonly use address and landmark references. On the base map, the definition and format of the address elements, spelling of the street name, definition of the landmarks and overall quality of the data are critical to successfully using attributes for location referencing.

A best practice for selecting a location referencing method for use in an application or for associating transit features to the base map is to:

- Use the most appropriate location reference for that application, and
- *Also* ensure that the reference method can be translated into the reference methods used by other applications. For example, if your paratransit application needs landmarks, use landmark references, but ensure that your base map also locates those landmarks using latitude-longitude, linear references, map grids, or whatever other reference method your other applications need, so that information can be easily shared.

Typically, there is a need to use multiple referencing systems to describe the location of a transit feature. (A list of the data descriptions supported by TCIP is included in Appendix C. Most commercial GIS programs include the algorithms to help you perform these transformations. Transforming between methods can be made easier by using a Location Table as described in Section 6.4.5.

### 6.3 Transit Feature Definitions and Attributes

Transit feature sharing is made easier by two best practices:

- Develop a consistent set of feature descriptions (e.g., data dictionary) that are stored and supported throughout your enterprise.
- Identify robust location referencing methods and the transformations between them for each feature.

This section describes key transit feature descriptions, their key attributes, and standard ways of assigning location reference methods to support these “best practices.”

### **Transit Feature Building Blocks**

There are only a few transit features that need to directly reference the base map. These building blocks are used to describe more complex transit features. The basic transit features include:

- Public Transportation Stops
- Time Points
- Patterns (or Time Point Intervals)

In the feature definitions, these three key features will appear as location referencing methods for more complex transit features such as route and vehicle assignment (block). They can be used to define the transit network or used in analyses and maps to help represent the spatial components of transit's business. A more detailed description of the importance of these building blocks is provided in Section 6.4.1.

### **Transit Features Categorized by Spatial Feature Classes**

These and other key transit features are described below in four tables that represent the major categories of transit spatial features. The transit features in the four tables can be further simplified into the three basic categories of spatial features that compose a base map in a GIS—points, lines and polygons. By dividing up the transit features in this way, there is a direct correlation between the transit feature and the base map construct. A point may exist as an event of a road segment or reference to a node on the base map. A line may be associated directly to a road segment (or link) of the base map or as an ordered series of points. In some cases, a transit feature such as “Station” may appear as a point in a small-scale map (e.g., 1:1,000,000) and as a polygon in a larger scale map (e.g., 1:25,000), for example as a bus terminal with multiple bus stops).

A number of the more complex transit features listed, such as Routes and Vehicle Assignments, are created from the three basic transit features: Public Transportation Stops, Time Points and Patterns. They are in a special category of transit spatial features called “Path.” Path” spatial feature class is defined in the TCIP standard (NTCIP 1405:2000).

The transit feature tables in the next three subsections list the transit feature name, a definition of the feature and common examples of key attributes of the transit field. The tables provide standardized definitions from TCIP or the Geospatial One-Stop Transit Layer (hereafter called One Stop-Transit). The Geospatial One-Stop is a Federal government program to develop “building blocks for a National Geographic Information Network.” See Sections 6.5 Resources for more information on the One Stop Project.

The “Description of Key attributes” column is not comprehensive, rather it contains key fields needed to support location referencing and fields common to many applications and business needs. [In the “Description of Key Attributes”, the location referencing methods are *italicized*.]

Following each table is a list of some location referencing methods that may be used to associate the spatial feature to the base map, facilitate data sharing or enable analyses. (A list of location referencing methods supported by TCIP is contained in parentheses in the section listing location

referencing methods following each table; the descriptions of the TCIP messages can be found in Appendix C.)

### 6.3.1 Point Features

**Table 6-1. Transit Point Feature Definitions and Attributes**

Point Feature Name	Definition (from One-Stop-Transit and TCIP)	Description of Key Attributes/Fields
<b>Public Transportation Stop Point</b>	An established location where public transport customers may board or alight from a transit vehicle in revenue service.	<ul style="list-style-type: none"> <li>• Identifier (unique id)</li> <li>• Location (latitude/longitude)</li> <li>• Relative location (address or intersection offset)</li> <li>• Amenity</li> <li>• Status (planned, active, under construction, retired)</li> <li>• Activation date/deactivation date</li> </ul>
<b>Time Point</b>	A point along a route where trips are assigned arrival or departure times.	<ul style="list-style-type: none"> <li>• Identifier (unique id)</li> <li>• <i>Location (latitude/longitude)</i></li> <li>• <i>Relative location (address; intersection offset)</i></li> <li>• Type [may depend on work rules]</li> </ul>
<b>Transfer Point (also called Transfer Cluster)</b>	A collection of one or more public transportation stops where transfer between routes is accessible and convenient.	<ul style="list-style-type: none"> <li>• Identifier (unique id)</li> <li>• <i>Associated Stop Points</i></li> <li>• Access description</li> </ul>
<b>Amenity</b>	The elements of a physical feature, a fixed location, or a transit facility. The amenities of a public transportation stop, for example, may include the shelter, platform announcement panel, and benches. An amenity may be described by one or more characteristics, or attributes, such as the year of construction or its current condition.	<ul style="list-style-type: none"> <li>• Identifier (unique id)</li> <li>• <i>Associated Stop Point</i></li> <li>• Bench</li> <li>• Shelter</li> <li>• Sign</li> </ul>

Point Feature Name	Definition (from One-Stop-Transit and TCIP)	Description of Key Attributes/Fields
<b>Public Transportation Vehicle</b>	A revenue conveyance in a transit fleet.	<ul style="list-style-type: none"> <li>• Identifier (unique id)</li> <li>• VIN</li> <li>• Vehicle capacity (seating, standing, wheelchair)</li> <li>• Mode</li> <li>• Attributes</li> <li>• Type</li> <li>• Manufacturer/Make/Model</li> <li>• Service Records</li> <li>• On-board components</li> <li>• Real-Time: <i>Current position (time/ location /heading)</i></li> <li>• On-route: <i>Associated pattern</i></li> </ul>
<b>Scheduled Event [NTCIP 1404: SchEvent]</b>	A transit message [or trigger] that is activated at a specified time, location, or both. Not to be confused with linear and point events of the Road model [base map], which may indicate the location of a subordinate transportation element or characteristic.	<p>Most scheduled events are associated with patterns. Events include:</p> <ul style="list-style-type: none"> <li>• <i>Associated pattern</i></li> <li>• <i>Associated stop (if appropriate)</i></li> <li>• Sign change (destination and interior)</li> <li>• Announcement</li> <li>• Fare Zone change</li> <li>• Driver message</li> <li>• Signal priority</li> </ul>

#### Point Location Reference Methods

Examples of point location referencing methods that might describe the location of the “point” transit features are:

- A universal reference such as State Plane [northing or easting], or Latitude and Longitude based on a declared datum (e.g., NAD '83, WGS '84). [TCIP: SpGeopoint, SpGeoLpoint, SpGeoDynamicpoint, SPpoint]
- A linear reference such as offset from intersection or linear distance from beginning of pattern, [TCIP: SpNodeOffsetpoint, SpNodePercentOffsetpoint, SpGeoOffsetpoint, SpMilepostpoint, SpNodepoint, SpLandmarkpoint] and
- [if available] an attribute such as address [TCIP: SpAddresspoint, SpIntpoint, SpIntOffsetpoint]

#### Best Practices related to Point Location Referencing Methods:

- Choose a method for each type of feature, use it consistently, and ensure that it can be translated to location referencing methods used by other applications in your agency.
- When using a commercial GIS, let the GIS transform between location referencing methods, but spot check for accuracy.

- Maintain an updated, spot-verified location referencing authority table (see Location Table discussion 6.4.5), rather than doing transformations “on the fly.”

### Amenities

Transit agencies treat or store amenities, such as bus stop amenities, in a number of different ways. Some include amenities as attributes within the public transportation stop entity (database tables), others create separate entities that relate (i.e., via a foreign key) the amenity to the public transportation stop. Typically, the amenity entity does not include latitude and longitude fields, rather it uses the stop identifier and an offset from the marker or curb edge as the location reference.

### ADA Attributes and Accessibility

Many transit agencies include non-descriptive fields to identify accessibility and conformance to Americans with Disabilities Act (ADA) requirements. These may not be sufficient, as many organizations have discovered. Appendix E contains a decision tree developed by Miami-Dade Transit to calculate conformance and the rules (metadata contained in the ADA compliance table) that support the decision tree. Based on the measurements collected through their field survey, they can calculate and populate compliance for each bus stop. The surveyed ADA measurements are related to the bus stop through the bus stop identifier.

## **6.3.2 Line and Path Features**

**Table 6-2. Transit Line Feature Definitions and Attributes**

<b>Line Feature Name</b>	<b>Definition (from One-Stop-Transit and TCIP)</b>	<b>Description of Key Attributes/Fields</b>
<b>Time Point Interval</b> (used by only a few transit agencies)	A one-way path of travel between two consecutive time points in a (transit) block.	<ul style="list-style-type: none"> <li>• Identifier (unique id)</li> <li>• Direction of travel</li> <li>• <i>Time points for termini</i></li> <li>• <i>Geographic path</i></li> <li>• Average distance a vehicle travels over the path (not centerline distance)</li> <li>• Type: revenue/non-revenue</li> </ul>

### Line Location Reference Methods

- Link or edge identifier from the base map [TCIP: SpLinkline]
  - Association with a specific base map may result in an inability to transfer the transit feature segment to another base map. Again, this reinforces the need for a common Location Table.
- Sequence of geographic points [TCIP: SpGeoline, SpGeoLline, SpMilepostline, SpSPline]



- This technique is typically used by on-board applications that use GPS or to generate a line or path on the fly using specific GIS functions (not all GIS tools support this function).

**Table 6-3. Transit Path Feature Definitions and Attributes**

<b>Path Feature Name</b>	<b>Definition (from One-Stop-Transit and TCIP)</b>	<b>Description of Key Attributes/Fields</b>
<b>Pattern</b>	A unique, non-branching, ordered sequence of time points, street links, or public transportation stops to be followed by a transit vehicle in scheduled service.	<ul style="list-style-type: none"> <li>• Identifier (unique id)</li> <li>• Direction of travel</li> <li>• <i>Time points for termini</i></li> <li>• <i>Geographic path (see location referencing methods below)</i></li> <li>• Average distance a vehicle travels over the path (not centerline distance)</li> <li>• <i>Ordered set of time points or time point intervals</i></li> <li>• <i>Ordered set of Stop Points</i></li> <li>• <i>Ordered set of Scheduled Event (trigger) points (variable record)</i></li> <li>• Type: revenue/nonrevenue</li> <li>• Timetable version</li> </ul>
<b>Trip (Temporal)</b>	A one way scheduled movement of a transit vehicle between starting and ending time points. A revenue-service trip will be an instance of a pattern.	<ul style="list-style-type: none"> <li>• Identifier (unique id)</li> <li>• Day type</li> <li>• Service type</li> <li>• <i>Associated pattern</i></li> <li>• Trip type</li> <li>• Associated routes, blocks, runs, scheduled events (optional)</li> <li>• <i>Ordered sequence of trip time points</i></li> <li>• Time point id</li> <li>• Arrival time</li> <li>• Time point attributes</li> <li>• Timetable version</li> </ul>
<b>Route</b>	A collection of patterns in revenue service with a common identifier.	<ul style="list-style-type: none"> <li>• Identifier (unique id)</li> <li>• <i>Patterns that apply (by day type and/or service type)</i></li> <li>• Trips that apply (by pattern, ordered by time)</li> <li>• Timetable version</li> </ul>

Path Feature Name	Definition (from One-Stop-Transit and TCIP)	Description of Key Attributes/Fields
<b>Vehicle Assignment (Block)</b>	A block is a sequence of revenue and non-revenue trips to which a transit vehicle may be assigned. The vehicle assignment begins when the vehicle leaves a vehicle base (pullout) and ends when it returns to a vehicle base (pull-in).	<ul style="list-style-type: none"> <li>• Identifier (unique id)</li> <li>• <i>Ordered sequence of trips</i></li> <li>• Day type</li> <li>• Pull-in/out time and location</li> <li>• Vehicle type</li> <li>• Timetable version</li> </ul>

#### Path Location Reference Methods

- A series of road or transit-domain lines (e.g., time point interval, pattern) that may be used to generate the path of transit feature [TCIP: SpL\_GeoRoute, SpL\_Linkroute]
- A series of linear references and the name of the link (e.g., street name) and its distance between points. [TCIP: SpP\_Addressroute, SpP\_Inroute, SpP\_IntOffsetroute, SpP\_Milepostroute, SpP\_Noderoute, SpP\_SProute]
- By using a transit feature building block, complex features such as the trip, vehicle assignment (block) and route only depend on a pattern (or TPI), time point and/or public transportation stop and not the base map. The GIS product that you use will map these transit features using the underlying spatial features.

### **6.3.3 Polygon Features**

**Table 6-4 Transit Polygon Feature Definitions and Attributes**

Polygon Feature Name	Definition (from TCIP)	Description of key attributes/ fields
<b>Facility</b>	Describes a facility or property owned by a transit property.	<ul style="list-style-type: none"> <li>• Identifier (unique id)</li> <li>• <i>Facility footprint</i> (polygon)</li> <li>• Facility name</li> </ul>

<b>Polygon Feature Name</b>	<b>Definition (from TCIP)</b>	<b>Description of key attributes/ fields</b>
<b>Station</b>	Same as Stop Point when the map resolution reduces the Station detail (e.g., start and end information) to a point on the map.	<ul style="list-style-type: none"> <li>• Identifier (unique id)</li> <li>• <i>Station footprint (polygon)</i></li> <li>• <i>Reference to stop point location (usually a centroid)</i></li> <li>• <i>Relative location (on-street/offset from street; address)</i></li> <li>• Amenity</li> <li>• Length/width of station</li> <li>• Portal (exit and entrance, if different)</li> </ul>
<b>Fare Zone</b>	The definition of location wherein a specified fare is associated. (The fare may be based on trips between two fare zones or as a fixed value at stop points contained in the zone.)	<ul style="list-style-type: none"> <li>• Identifier (unique id)</li> <li>• <i>Fare Zone coverage (polygon)</i></li> </ul>
<b>Radio Zone</b>	The coverage area over which a radio transmitter operates.	<ul style="list-style-type: none"> <li>• Identifier (unique id)</li> <li>• Channel list</li> <li>• <i>Zone coverage (polygon)</i></li> </ul>

Location referencing methods for polygons are typically inherent in the GIS; the GIS tool may build a table based on geographic points (latitude/longitude), or nodes (vertices) and links (edges) from the base map. Generation of the polygon may be “on-the-fly” or stored as a sequence of nodes/links in a file. “Buffers” around a point (e.g., bus stop zone) or line (e.g., walking distance for ADA eligibility) are typically generated on the fly by an internal algorithm in the GIS or spatial analysis software.

TCIP has standardized different ways of describing the location reference method for the polygon including:

- Centroid (SpCentroidpolygon)
- Geo points (SpGeopolygon, SpP\_Sppolygon – state plane)
- Other types of points (SpP\_Intpolygon, SpP\_Nodepolygon)
- Links (SpL\_Linkpolygon)
- Other types of lines (SpL\_Geopolygon, SpL\_AddressRangepolygon)

#### **6.3.4      *Relationships Among Transit Features***

Relationships among transit features are very complex. There are many-to-many relationships among stops, routes, patterns and blocks. For example, many stops are in a route, and a stop may be used in many routes. Similarly, a route consists of multiple patterns.

In order to aggregate transit features and data correctly, the relationships between transit features must be carefully defined and implemented. If they're not, location-referencing errors can be introduced, ITS applications can operate incorrectly and valid analyses may not be possible. The tables in Section 6.3 define the basic relationships among the features. Each transit agency may have additional complex relationships that need careful definition.

The transit industry-wide effort to develop the TCIP definitions made significant in-roads in defining the transit features and relationships. The relationship among TCIP Schedule objects, which include many of the complex service transit features, is illustrated in Figure 6-1 below.

Clearly, doing every complex relationship/association as link attributes is not the most elegant practice. However, it has the advantages of being easily understood by GIS operators, and it works well with the underlying relational database structure of the GIS package. For less complex applications, non-customized applications, and those using off-the-shelf GIS, this may well be the best practice. Since many location referencing schemes are based on links, it also makes translating between location referencing methods relatively straightforward.





### 6.3.5 *Associating Performance Measures with Transit Features*

Although not directly related to location referencing, associating performance measures as attributes to transit features results in some of the greatest benefits that may be achieved by “mapping” transit features. A GIS may be used to perform innovative analyses and to help transit visualize service quality, if performance measures are linked to spatial features and locations are referenced appropriately.

Table 6-5 provides examples of how performance measures can be associated with transit features. The raw data may be attributed to transit features such as time points, stop points, patterns, time point intervals or trips. Any of these attributes may also be categorized by temporal characteristics of a transit feature such as day type, time period or time of day.

**Table 6-5: Transit Feature Attributes: Performance Measure**

Transit Feature	Performance Measure
Time Point	<ul style="list-style-type: none"><li>• Peak load</li><li>• Schedule adherence (e.g., status—early, late; or on-time)</li></ul>
Stop Point	<ul style="list-style-type: none"><li>• Schedule adherence (e.g., status—early, late; or on-time)</li><li>• Boarding and alighting counts</li><li>• Vehicle arrival and departure times (or dwell time)</li><li>• Vehicle load point</li><li>• Passenger wait time</li><li>• Real-time attributes<ul style="list-style-type: none"><li>– Estimated time of arrival</li><li>– Time left last stop</li></ul></li></ul>
Time Point Interval or Pattern (by trip or time of day)	<ul style="list-style-type: none"><li>• Running time</li><li>• Load</li><li>• Incident (e.g., aggregate by incident category)</li><li>• Schedule adherence (e.g., status—early, late; or on-time)</li></ul>

## 6.4 Recommended Practices for Transit Features

Transit features constitute the core elements of the transit network. Building the network requires that a robust transit feature infrastructure be developed. There are three key goals in building the transit network:

- Clearly describe the transit features and their relationship to the base map
- Clearly describe the relationship among transit features
- Support the ability to unambiguously transform a feature from one location referencing method to another

Six important steps for accomplishing these goals are as follows:

- Identify the Transit Feature Building Blocks
- Describe Transit Features relative to these building blocks

- Build a Data Model
- Manage Location Referencing Methods
- Build a Location Table
- Store the Transit Features Effectively

These steps are recommended as best practices and are outlined in the sections below. Some of the step's elements were referenced earlier in this section.

#### **6.4.1      *Identifying Transit Feature Building Blocks and Associating Them With the Base Map***

The first step in associating the transit features to the base map is to identify the fundamental transit features or building blocks that will be used to construct the transit network. Those features and the network will serve as a framework for aggregating and disaggregating transit data for analysis, display and use in ITS applications.

As stated before, transit features may be described as points, lines/paths or polygons. Points and lines are necessary to define the transit network and form linkages to the base map. Polygons can be developed later from the points and lines, so they are not considered fundamental building blocks.

Table 6-6 lists the basic transit features, or building blocks, that are necessary for developing more complex transit features.

**Table 6-6: Basic Transit Feature Definitions**

<b>Transit Feature Name</b>	<b>Definition</b>	<b>Source</b>
Public Transportation Stop	An established point where public transportation customers may board or alight from a transit vehicle in revenue service.	Geospatial One-Stop (GOS) Transit
Time Point	A location along a pattern where trips are assigned arrival or departure times	Revised GOS Transit
Pattern	A unique, non-branching, ordered sequence of time points, street links, or public transportation stops to be followed by a transit vehicle in scheduled service	GOS Transit

[Note on Trip Segment or Time Point Interval (TPI). TPIs or trip segments are used as a more basic construct to build patterns when complex route structures, such as branching, short turn trips and alternative paths and stop configurations are used by schedulers. A trip segment is illustrated in Figure 6-2 and the relationship of the to other transit building block features are illustrated in Figure 6-1.]

As a best practice, an agency inventory of these three feature sets is necessary before other, more complex features are developed.



### Field Surveys of Building Block Features

When these building blocks are used for on-board applications, field surveys of the building blocks are recommended. The field surveys can collect the absolute locations (latitude/longitude, and altitude or level) and as-driven distances (along a pattern) that are needed when the scale or accuracy of a base map is inadequate. For an accurate Public Transportation Stops inventory, a field survey is almost always required to measure the location and length of the stop. ADA characteristics such as width of access, distance from the curb, qualities of the curb cut are usually collected during the field survey as well. Time point locations and pattern paths may also require field surveys, particularly for schedule adherence and positioning of key triggers along a vehicle assignment. The field survey should not only collect absolute location information, but also information that will support its association with the linear network of the base map. These best practices imply the use of certain location referencing methods, typically latitude and longitude, and a linear referencing method or attribute reference such as offset from an intersection or address.

The transit feature may be described as being an (spatial) event on the base map. The best practices for referencing these features to the base map are listed in Table 6-7.

**Table 6-7: Association of Basic Transit Features with Base Map Features**

Transit Feature Name	Base Map Feature
Public Transportation Stop	<ul style="list-style-type: none"><li>• Geographic point (lat/long) <i>and</i></li><li>• Reference to the road network (e.g., offset from at and cross streets with side of street, or address)</li></ul>
Time Point	<ul style="list-style-type: none"><li>• Geographic point <i>and</i></li><li>• Intersection (at and cross streets)</li><li>• [Offset from Public Transportation Stop]</li></ul>
Pattern	<ul style="list-style-type: none"><li>• Ordered sequence of directed Road Segments (tied to base map);</li><li>• Pattern should terminate at both ends with a time point; and</li><li>• [As-driven] distance between time points and/or distance between public transportation stops</li></ul>

Appendix D shows how the GOS model associates basic transit feature building blocks to base map features, such as point, line, polygon, using multiple referencing techniques. The model incorporates the location referencing requirements described in Table 6-7.

### Issues Related to Mapping Field Survey Data to Base Map

Linking survey or navigation sensor (AVL) data to a base map requires a few transformations. Each transformation introduces “errors” into the measurement. As such, additional steps may need to be executed to correct or realign the locations on the base map.

1. Transform the collected data to the same projection as the base map.

- Step one ensures that the projection of the spherical coordinates (e.g., latitude/longitude) is the same.
- 2. Position [projected] surveyed location to the relative position, e.g., distance along the pattern or from pullout (for AVL tracking), or position relative to intersection (for field survey).
  - Step two ensures that the position is aligned to achieve the proper relation to the transportation network.
- 3. The original surveyed and tracked position should be archived or associated with the transit feature.
  - Step three ensures that the absolute or universal position is kept for reference purposes.

#### **6.4.2      *Describing Transit Features***

The definitions of the core transit features for a transit agency should be based on the business requirements of the organization. Potential stakeholders and systems that might use transit spatial data should be identified. After identifying the potential data uses, features and attributes needed, core transit features can be defined and an inventory developed. Specifically, the steps for describing the transit features are:

- Identify core data and define them across the enterprise (including their attributes)
  - A list of core transit features and key attributes are listed in Section 6.3.
  - A list of transit application and functions that use spatial data are contained in Appendix G.
- Identify location referencing and indexing requirements
  - Identify definition and usage of unique identifiers for each feature. The identifier must be able to serve multiple applications; and, be flexible and large enough to accommodate growth (e.g., in systems, number of facility locations, fleet, etc.). Often it must be able to handle “retiring” an ID number rather than reusing it, such as with Bus Stop identifiers. (Procedures for retiring the bus stop identifier should deal with when is it appropriate. For example, what if the stop is moved 10 feet on the same side of the street? You may find it useful to describe in a document the changes that warrant retiring the ID.)
  - Identify the types of location referencing methods required by applications and systems that use the feature.
- Identify data integration issues related to core datasets. Typically, integrating data requires human intervention and judgment. For example, building a transit pattern is not easy to fully automate, when it includes an ordered set of stops, annunciator/sign triggers, and the correct assignment of stop announcements and sign messages. Once the data relationships are established, many data sharing activities may be facilitated by a robust Location Table or other transformation techniques.

Some basic best practices based on input from transit staff include:

- Gather location referencing and application data sharing requirements prior to performing field surveys, particularly for a Bus Stop Inventory (See recommendations in the Bus Stop Inventory Guidebook that is listed in Section 6.5, Resources.)

- Include key location referencing methods in the feature description or in the Location Table that aggregates all the applicable location referencing methods.
- Use unique ID numbers for the features; do not reuse an ID number more than once in the life of the inventory. For example, if a stop is retired, do not recycle the Stop ID number.
- Understand the maintenance costs and resources related to updating the dataset.

### **Lesson Learned**

Inconsistencies between the data description assumptions and the “real world operations” can still occur. A difficult to detect error can arise when schedulers link patterns, time points and bus stops to describe transit service. When defining patterns and building the blocks (linking trips), if the bus stop locations are not physically contained within the trip’s pattern, a scheduler may not realize that a bus stop exists on the far side of an intersection. As a result, the scheduler may allow the vehicle to turn without servicing a far-side stop. If this inconsistency is not caught, problems arise for customers and data analyses. A customer information system may incorrectly assume that passengers can alight at the stop location and APC applications will produce inaccurate ridership data at a route or trip level. Eventually, the mistake will be found, but not without cost.

A data model helps uncover these relationships by explicitly drawing attention to the specific associations. For that reason, it is helpful to illustrate what is meant by pattern and disseminate that drawing throughout the organization. A GIS group may find it useful to define a “trip segment” [as described by Orange County Transportation Authority (OCTA)] or “time point interval” [as described by King County Metro]. Figure 6-2 illustrates how the published route map relates to time points, bus stops, time segments and patterns. Together these patterns compose a single route direction.



76

### **6.4.3      *Building a Data Model***

A robust data model that defines the relationships between transit features can support analysis, reporting and ITS needs. Due to differences in organizational structures, work policies, and cultures at various agencies, defining a single relational data model for all of transit is very challenging. A data model is often helpful for defining how applications will share information, and how data will be linked for analyses and reporting needs.

Developing a comprehensive data model that supports all of a transit agency's data is a daunting task. Few organizations have been successful when they try to develop and implement a comprehensive model. Limited data models that represent parts of the enterprise linked by key entities is a "modular" approach that has been successful. This approach assumes an organization has a good idea about the key relationships among its data. Issues related to pieces of the model are analyzed and developed in an evolutionary approach. Using a reference model such as one developed by another transit agency, GOS (limited to geospatial data), or one developed by the Europeans (i.e., TRANSMODEL) will help mitigate the risk of this approach. Section 6.5, Resources, includes references to GOS and TRANSMODEL.) Other ways to limit the impact of changes to your data model are discussed in Sections 3.7 and 9.4

One of the major advantages of developing a data model is the implementation of internal referential integrity procedures. These procedures perform format checking, verification of identifier uniqueness, business rule checking (e.g., "patterns must terminate with valid time points"), and updating/flagging of related features (e.g., if a bus stop is retired, the pattern that contains an ordered sequence with that bus stop is identified as changed, too). The integrity checking may apply to both attribute and spatial relationships of the transit feature. These internal consistency rules can help ensure that the location referencing attributes are accurate and translate appropriately between applications.

### **6.4.4      *Considerations in Managing Location Referencing Methods***

There are strategies for managing the quality (i.e., accuracy, consistency, completeness) of the location referencing methods. Human and system errors will occur when interpreting location references or converting between location referencing methods. Steps may be taken to reduce errors in translation among location referencing methods across the enterprise such as:

- Implement address matching standards (standard formats for storing addresses)
  - Standards exist which describe various ways for formatting an address or address range, see Section 6.5, Resources. As a best practice, consider using the same format as the U.S. Postal Service or the Census Bureau. Many commercial GIS packages will allow you to pick from several standardized formats. The most important best practice is to be CONSISTENT across all applications within an agency.
- Develop a master street name alias file
  - GIS vendors may use different formats to store street name alias files.
- Build a Location Table (see discussion in Section 6.4.5). Even if your applications are relatively simple and your agency small, building a Location Table to serve as an "authority" on location references is a very important best practice.

- Develop internal procedures for associating surveyed measures of locations (such as GPS) to the base map (using transit features and attributes, or street centerline attributes). Many commercial GIS packages have tools to help you do this.

For procurements that involve existing transit-managed datasets, a document that details the data quality and location referencing methods required to access the data may be helpful.

#### **6.4.5      *Building a Location Table***

A location referencing authority table (Location Table) defines the authority for indexing location references. The table serves as the “authority” for location references for transit point features. It can also serve as the authority for looking up transformations among location referencing methods. For example, a bus stop would be identified in the table with a unique ID number and might be associated with several location reference methods such as the latitude and longitude, offset from the “at” and “cross” streets, and/or address.

**Benefits of a Location Table.** Having all of the spatial point data in a single table has several advantages. First, data maintenance is greatly simplified when the location reference methods of multiple features only have to be updated in a single table (e.g., amenities at a bus stop). Tasks like projecting to a new coordinate system, re-associating points with jurisdictional or zip code boundaries can be done in one place. These changes are automatically reflected in all associated datasets by virtue of the referential association.

Another key advantage is the ability to remove the spatial requirement from other independent applications while linking them back to a robust and comprehensive spatial description. Applications that are largely tabular in nature do not need to have the added requirement of spatial tools and their associated licensing fees, yet can still be tied back to the spatial world through a simple location ID reference. For example, a system to track graffiti on shelters can easily be given a simple drop-down window to find location ID numbers of shelters when completing graffiti report forms. Analyses can be completed without spatial tools to report statistics by jurisdiction, zip code or route. At the same time, spatial analyses can easily be accomplished through the location ID link.

**Best Practices.** Typical best practice for assigning location references to a transit feature assumes that at least two or three types of location reference methods are included in the feature record. For example, the bus stop inventory requires multiple location references to meet the needs of its users. This practice associates the transit feature with multiple location referencing methods, e.g., geographic (position on earth), linear (position relative to the transportation network), attribute (such as address, position relative to human interpretation).

Some agencies take this recommendation to the next level by creating a special table that aggregates location referencing methods for each feature. A Location Table serves as the *authority* for transformations among location referencing methods. All the methods are generated once, migrated to applications once, and managed in a single location (the Location Table). This approach reduces the chance of introducing errors (due to transformation processes), reduces the need to cascade changes to application datasets when changes to a base

map are introduced, and provides a single point of reference for managing changes to features associated with a base map. This particular best practice applies regardless of whether your agency is purchasing an attributed transit database or doing the feature attribution in-house. This single best practice alone can significantly reduce errors in location referencing.

### **Location Table Best Practice Example (Tri-Met)**

A best practice method for building a Location Table is described below as it is implemented at Tri-Met.

**Location ID.** The location referencing authority table is a centralized Location Table that stores all point data referenced within the transit enterprise data model. An arbitrary numeric identifier, known as the Location ID, is assigned to each new location point of a transit feature as it is sequentially introduced to the database.

**Location Table Attributes.** All descriptive attributes of the location itself, including the spatial attributes, are kept in the Location Table. The descriptive attributes of the Location ID are confined to those that are location based. Descriptive attributes include items such as “at” street, cross street, public description, and direction of travel. Examples of spatial attributes include latitude, longitude, state plane x and y coordinates, zip code, fare zone, and jurisdiction.

Certain spatial fields are included in both the feature record as well as the Location Table. Values such as zip code, normally obtained through “point in polygon” analysis in the spatial world, are made available to the non-spatial applications through relational queries. Tri-Met generates line and polygon features from point information in their GIS.

Attributes such as number of shelters or type of location (e.g., stop or time point) are more appropriately maintained through other relational tables sharing the common Location ID reference.

**Managing and Maintaining Fields in a Location Table.** In many applications, the currency of the data, indicated by a date/time tag, must be understood and managed carefully to avoid problems. By referencing the Location Table using the Location ID, the spatial data is essentially removed from the primary dataset. This introduces the danger of an unintended ripple effect when a key attribute in the location table is updated. It is important to maintain consistency of the reference across time. The Location Table cannot lose the old record keyed to the Location ID; therefore, a method to manage the updated records must be implemented.

Adding one attribute to the unique identifier (primary key) of the Location Table, called `location_end_date`, can achieve this consistency. When a location ID is first introduced, the `location_begin_date` is set to the date of first use (see Example Table #1 below: 5-1-03 is the new date). The `location_end_date` is set to a date in the far-off future (in Example Table #1, `location_end_date` is 12-31-9999). If a change is made to one of the Location Table attributes, for example, if the `Public_Description` is changed, the `location_end_date` for the existing row is set to the date of closure (in Example Table #2, the date is now 6-30-03). In addition, a new row (or record) may be introduced for the Location ID. On the new record, the `location_begin_date` value is set for the original `location_end_date` plus one day later (7-01-03, and the

location\_end\_date is set to a date in the far-off future (12-31-9999). Using this approach, a consistent set of values is provided for the Location ID with the referring transit feature matching the Location ID, and inclusive of location\_begin\_date and location\_end\_date.

Example of the Public\_Description changing:

**Example Table #1**

Loc_ID	location_begin_date	location_end_date	Public_Description
5	5-1-03	12-31-9999	Union & 5th

**Example Table #2**

Loc_ID	location_begin_date	location_end_date	Public_Description
5	5-1-03	6-30-03	Union & 5th
5	7-1-03	12-31-9999	MLK& 5th

In this manner, historical data can be matched accurately. Alternatively, if the location changes (e.g., nearside to farside), a new ID (and new record) is created and the old ID is retired.

[Note: An alternative approach is presented in the Section 6.5 Resources, Time-Tagging Database Strategies.]

#### **6.4.6 Storing Transit Features**

Transit features—points, lines or polygons, must be maintained in a GIS to ensure accuracy and to generate topology. A GIS allows these features to be exported in a variety of different formats, making the data easily accessible to other GIS, computer aided design, and graphics applications. There are some restrictions when exporting route topology.

With recent technological advancements, it is now possible to store and maintain your spatial data directly in a relational database. There are several advantages to this, including the ability to maintain both spatial and non-spatial data in the same database, in a consistent way, with standardized associations.

When using commercial GIS products, the vendor's gateway tool is needed to access spatial data layers stored in relational database through a GIS. If however, you have non-GIS clients accessing the data, you will need to implement the spatial component available with your commercial database. These products can be utilized together so all clients can access the data and perform spatial queries. Every major GIS vendor offers gateway tools and most commercial database vendors provide some support for accessing spatial data.

Transit features that are defined as "point features" can be stored most efficiently in a relational database table, with the location of the features stored in a related Location Table. For example, the attribute table for bus stops would be linked to the Location Table by the unique identifier, Location ID. By storing the data in this manner, the data can be maintained by means other than a GIS, such as an upload from a GPS unit. It can be accessed by non-GIS clients, and accessed by GIS clients by dynamically displaying them as "XY Events". The advantages to this



approach include the ability to access updated information immediately, as it is being maintained. Spatial indexes that are required for the GIS gateway do not need to be generated after updates.

Transit features that are defined as “polygon features”, such as fare zones or incident response dispatch areas, must be maintained in a GIS to build topology. They can be accessed by both GIS and non-GIS clients.

Routes can be generated directly from the base map or centerline file by various methods:

1. Routes can be digitized graphically; however, use of the information will be very limited. This method does not allow temporal information, such as schedule information for each trip by bus stop, to be associated with the route. In order to associate the temporal data, route topology must be generated at the pattern level and associated with stop and schedule information.
2. The location of the stops can be used in combination with schedule data to generate a route based on the shortest path. The advantage of this method is that the unique ID numbers of the street segments do not need to be stable. It can be relatively easy, in theory, to regenerate routes using an updated base map or centerline file, assuming the line work is complete, base map attributes (such as bus only lanes, reversible lanes, road restrictions) are incorporated into the route generation algorithm, and the stop locations can be re-projected. Automated generation of the path works best for routes with a concentration of bus stops along the path of travel. For express routes where the drivers are expected to travel along a certain path, this approach may help derive the path, but will still need editing by a knowledgeable technician. Regardless, *automated generation of routes needs to be checked for accuracy and completeness.*
3. The routes can be generated from a table containing the ordered sequence of street segment ID numbers. This method may require extra work if the spatial data is being referenced onto a different base map. If it is, sometimes the data needs to be conflated to transfer the street segment ID numbers.

## 6.5 Resources

### Bus Stop Inventory

“Bus Stop Inventory Best Practices and Recommended Procedures”. Bus Stop Inventory Task Force Transit Standards Consortium, September 19, 2001.

To obtain a copy of the Guidebook, contact:

Transit Standards Consortium

91 West Clinton Ave., Suite 2

Tenafly, NJ 07670

Telephone: 201-567-0088

Fax: 201-567-8066

<http://www.tsconsortium.org/TSC2003/BSIOrders/BSIOrderForm.htm>

### Geospatial One-Stop

The Geospatial One-Stop is a Federal government program to develop “building blocks for a National Geographic Information Network. It is one of 24 e-government initiatives sponsored by the Federal Office of Management and Budget to enhance government efficiency and improve citizen services. Geospatial One-Stop will make it easier, faster and less expensive for all levels of government and the public to access geospatial information.

**“Geospatial One-Stop will provide tools for all levels of government—local, State and Federal—to work together more effectively to:**

- Facilitate sharing of geospatial information
- Improve planning for future investments in geospatial data
- Expand collaborative partnerships that help leverage investments and reduce duplication
- Collaborate on the development and implementation of standards that encourage sharing and use of best practices and advance implementation of the National Spatial Data Infrastructure.”

<http://www.geo-one-stop.gov/about/index.html>

### TRANSMODEL: Transit Reference Model Standard (European)

Transmodel is conceptual data model that describes the functions and relationship among public transport entities. The model was developed by the European public transportation community as part of the European standards development organization, CEN.

<http://www.transmodel.org>

### Time-Tagging Database Strategies

Richard T. Snodgrass, Developing Time-Oriented Database Applications in SQL. Morgan Kaufmann Publishers, San Francisco CA, 2000.

## SECTION 7 SPATIAL DATA QUALITY

### 7.1 Overview

This Section will define data quality categories and describe how to develop a data quality process and infrastructure in your organization. In addition, it will discuss issues and approaches to documenting data quality in metadata.

Best practices for achieving spatial data quality include the following:

- Quality should be defined in terms of fitness for use for specific applications
- Quality should be documented with appropriate metadata, so that users within your agency can efficiently determine if a dataset is appropriate for an application
- Quality includes positional accuracy, attribute accuracy, logical consistency, currency and completeness
- Quality standards and procedures should be implemented both when acquiring datasets and during the operations and maintenance phases

Data quality is a deceptively simple concept, because many people think of data quality purely in terms of positional accuracy. In actuality, there are additional components of quality that should be considered. In addition, the tools used in collecting spatial data and the tools' capabilities should also be known, because they can establish a maximum possible accuracy.

For the transit professional, however, the relevant quality is the “fitness-for-use” of a particular piece of data for your application. The accuracy of the data with regards to its position on a mathematical model of the earth's surface (the reference ellipsoid) may not be the quality that is most relevant for transit scheduling applications.

If you ask yourself what you really need to know about the data in order to determine whether it is "good enough" for your application, positional accuracy and completeness should definitely be on your list. Your list would probably include whether the data are current and whether it is in a format compatible with other applications in your agency. You might ask whether the scale of the digital map source provides adequate data resolution and positional accuracy, whether the features are consistently defined across the dataset (and consistent with the way features are designed in your agency), and whether the coverage is roughly what you need. These all are important measures of quality.

To effectively determine the fitness for use of a particular dataset, you need to have detailed information about that dataset. Since you might want to analyze the data in the future, you need to maintain data about the data (metadata) in a format that is consistent across your organization and comprehensive in detail. Section 11 has more information about general metadata principles; this section discusses some of the most important metadata needed for effective data quality management.

### **7.1.1 Problem**

Problems will arise from data quality issues that adversely affect data analysis and reporting, service performance, customer information, and other aspects of meeting transit's mission.

General problems arise from:

- Inadequate or poorly communicated definitions of user requirements and data specifications
- Lack of quality checking procedures when acquiring or updating spatial data
- Failure to involve users in defining data testing conditions and error “traps” needed for catching both random and systematic quality problems
- Inadequate documentation on collection and maintenance procedures
- Multiple, inconsistent datasets of the same data (including various field collections of similar data, e.g., bus stop and route alignment)
- Inconsistent ways of referring to location (e.g., address formats and use of street aliases, use of different coordinate systems)

These problems result in obstacles such as:

- Limited ability to update or transfer data (transit features or transportation network) between applications
- Limited ability to integrate core transit data for advanced technology projects
- Inappropriate, inaccurate or misleading information for generating maps, reports, performance measures and for supporting Intelligent Transportation System (ITS) applications such as Automated Vehicle Location (AVL) or Trip Planning

### **7.1.2 Purpose**

This section will define data quality categories and describe how to develop a data quality process and infrastructure in your organization. In addition, it will discuss issues and approaches to documenting data quality in metadata, data about data.

Specifically, this section describes best practices related to:

- Defining a Quality Process
- Internal Data Policies and Standards
- Data Ownership
- Data Committees
- Developing a Data Dictionary
- Quality Checking

## **7.2 What is Quality?**

As stated in the Overview, quality is the fitness for use of a particular piece of information or dataset. To obtain quality, the elements of quality need to be understood so data checking can occur. In addition, quality is achieved and maintained by employing effective data management practices, determining fitness for use, and standard methods for data definition and attribution.

### **7.2.1      *Data Quality Principles and Data Management***

Data quality principles and procedures should drive the specific tools an agency chooses to share, integrate and use in spatial data applications. The first step in effective data management is identifying and articulating clearly what the agency expects to achieve with spatial data. However, many data users have a short attention span and limited technical grasp of data quality concepts and requirements.

Effective data management means applying general management best practices to spatial data management. In many organizations, an appointed data committee of end users is lead by an individual who does understand technical data quality concepts and requirements. The data committee establishes metrics for map performance that are easy-to-understand, handles the interface issues among applications, and coordinates maintenance and deployment schedules. The objective of this management function is to make the enterprise more efficient in its specification and use of data and transit functions. A data committee is not the only way to implement effective spatial data management, but it is a good tool and common “best practice.”

Spatial data management policies and procedures are also needed to facilitate regional data sharing and regional ITS applications. Transit agencies *use* spatial data from others such as vendors, other transportation agencies, Metropolitan Planning Organizations (MPOs), the tax assessor’s office and Departments of Transportation (DOTs). Transit agencies *send* data with spatial components to traffic control centers, planning organizations, DOTs and others. Finally, transit may also *operate* regional ITS systems that use spatial data such as regional Ridematch systems, Traveler Information systems and AVL. Metadata is also needed to ensure that the appropriate data are located and used.

In short, the enterprise data management policies and procedures are the “best practices” within an organization or a region for supporting the spatial data life cycle.

### **7.2.2      *Data Standards Help***

Data standards provide a framework for how data are managed in an organization and define corporate data policies (e.g., data dictionary, data custodians, data acquisition policies, core datasets, application interoperability procurement standards, etc.).

Enterprise and regional data management policies and procedures may also establish internal standards for tables, formats, software coding languages, database management systems, corporate data reference models, corporate base maps, street addressing standards, street alias look-up tables, location referencing systems (with alternate methods), data aggregation techniques and transit feature definitions, reporting/metadata standards and forms, data collection procedures and reporting requirements.

### **7.2.3      *Data Quality Elements and Determining Fitness for Use***

Data content and data quality needs depend mainly on application and user requirements, i.e., different applications need datasets with appropriate and necessary content and quality.

Understanding the quality elements used to report and evaluate data quality is key to procuring, operating and maintaining successful information-driven applications and to implementing effective quality checking procedures.

To find the appropriate dataset, documentation on the base map, other geographic data and transit features, metadata should provide answers to questions related to application and business needs reporting criteria. The set of questions below was developed by the Federal Geographic Data Committee (FGDC) to address “Content Standards” for documenting quality. Questions include:

1. “Is information available that allows a user to decide if the data are suitable for the purpose?”
2. What is the positional and attribute accuracy?
3. Are the data complete?
4. Was the consistency of the data verified?
5. What data were used to create the dataset?
6. What processes were applied to these sources?” [From *Content Standards for Digital Geospatial Metadata Workbook*. Version 2.0; FGDC, May 1, 2000]

These questions address five areas of quality that are documented in the FGDC Metadata standards and are supported by the Framework Data Content Standards program. The five data quality elements include:

- Attribute Accuracy
- Positional Accuracy
- Logical Consistency
- Completeness
- Lineage

These categories relate to both transit feature and base map data. They are explained in greater detail in Section 7.3 below.

### **7.3 Data Quality Elements**

Quality is defined as the measured assessment of data to a set of criteria. The criteria discussed in this section fall into the data quality element categories described above: attribute accuracy, positional accuracy, logical consistency, completeness, and lineage and currentness. Aspects of the major data quality elements or categories are defined by a number of international standards on geospatial datasets and may apply to the base map or to domain feature datasets (See Technical Memorandum #2 referenced in 7.5 Resources).

In some cases, measuring quality may be straightforward. For quality measures such as logical consistency, the metric for quality may be “pass/fail.” For example, if the quality measure is whether bus stops are associated with the correct side of the street in the base map, the base map is either correct or wrong. In other cases, the measure may be more subjective, such as judging how well exceptions are handled.

### **7.3.1      *Positional Accuracy***

Positional accuracy is the assessment of the closeness of a feature location to the actual position. Accuracy may be measured in different ways: absolute and relative. These measures are based on GIS/cartographic definitions for accuracy. *Absolute* accuracy may be defined as how close the measured position is to the true position relative to the location of a surveyed control point. State Plane coordinates are examples of absolute measurements. The accuracy of the absolute measurement of an object is described by the closeness of its State Plane position to the true State Plane position. *Relative* accuracy may be defined as the relation of a location to some other known position, e.g., address, offset from intersection centerline, distance from the start of a route. The accuracy of a bus stop location relative to the intersection corner is the measured distance to that corner. Both absolute and relative positional accuracies are needed for assessing whether the dataset is appropriate for a transit application or analysis.

Accuracy is typically described by the error propagated by the measuring instrument, data collection process, and various business processes to which the data is subject. For example, errors are introduced throughout the dataset life cycle, when it is acquired, transformed (recorded), managed, selected, integrated, derived and reported. Data may be subject to errors in the measurement precision (how many decimal points are valid); accuracy may be described as the spread of repeated measurements (e.g., standard deviation, root mean square, or covariance), at a later stage, a surveyed location may be snapped to road feature of the base map, and the distance calculated from a set of links between stop points. An understanding of which of these operations applied may reduce or enhance the accuracy of the data will provide insight into whether a dataset is fit for use or meets the requirements for an application.

### **7.3.2      *Attribute Accuracy***

Attribute accuracy is an assessment of whether required information about an attribute is correct. For example, are one-way roads correctly identified? Are roads correctly classified by the type code? Formally, attribute accuracy is the association and assessment of the accuracy of feature characteristics in the dataset, and the assignment of values to those characteristics.

Two key aspects define attribute accuracy, temporal accuracy and thematic accuracy. Temporal accuracy assures up-to-date information and the correct version. It pertains to temporal fields and temporal relationships of features such the date the data was collected in the field survey, the date of the last update, and the date the element was activated or deactivated. Section 6.4.5 discusses issues related to “time tagging” fields in databases. Documenting temporal accuracy is directly related to time tagging fields in the database. See resource on time tagging strategies in Section 6.5. Thematic accuracy defines quantitative fields, the correctness of non-quantitative attributes and correctness of the classifications of features and their relationships. Lighting area and brightness of coverage at a stop point, type of fixture, and its relationship to the stop point zone are examples of attributes associated with thematic accuracy.

The BUS STOP entity is an example of a transit feature that includes attributes. Associated with each field are rules for assessing attribute accuracy. In Table 7-1, a sample of BUS STOP attributes are shown as an example.

**Table 7-1: Sample Bus Stop Attributes**

<b>Inventory Element/Group</b>	<b>Inventory Data Elements</b>	<b>Description/Comment</b>
Bus Stop Identification		Unique integer for each bus stop that was ever used by the transit agency.
Location Coordinates	Latitude	Decimal degrees with a minimum of five decimal places (e.g. 44.12345)
	Longitude	Decimal degrees with a minimum of five decimal places (e.g. 114.12345) (longitudinal error varies by latitude)
	Altitude	[optional] Represent by 10 <sup>th</sup> of meters as xxx.x and specify whether it is an offset from street level or Mean Sea Level
Intersection Relative Coordinates	On-street	May represent street by segmenting <ul style="list-style-type: none"> <li>▪ pre-direction</li> <li>▪ prefix</li> <li>▪ street name</li> <li>▪ street type (suffix)</li> <li>▪ post direction</li> </ul> <i>(see Geographic Information Framework Data Content Standards For Transportation Networks: Transit)</i>
	At-street	See comments under on-street
	Distance from intersection	In meters to 10 <sup>th</sup> of meters
	Direction of route	Use NTCIP 1403 SCH-RouteDirectionName (north, south, east, west, southwest, southeast, northwest, northeast, inbound, outbound, circular, destination, clockwise, counterclockwise, name)
	Placement relative to intersection	Use NTCIP 1405 SP-RelativeDirection (nearside, farside, mid-block, at, between, farside-mid-block, nearside-mid-block, opposite)
Shelter Type		Matched to code list for available shelter types, including code for “none”

### 7.3.3 Logical Consistency

Logical consistency means that the dataset complies with valid business rules, permissible attribute values and relationships. The quality measure is the degree of adherence to logical rules of the data structure, attribution and relationships. The data structure can be conceptual, logical or physical. For example, stop point identifiers are unique and should not be recycled; street names must be validated against a master list of street names and alias table. For the base map, there are rules for verifying the correct number of vertices and edges (e.g., links and nodes), their topology and other logical constructs. Logical consistency relationships may be checked and managed automatically, by look up tables or by consistency and integrity rules enforced by procedures in the database management system.



For the bus stop example above, logical rules might include:

- The intersection referenced must use a standard naming convention and spelling; matching all street names against a street names and alias table will check for use of standard spelling
- The bus stop identifier must be a valid, unique bus stop ID
- All distances will be measured in linear meters (as opposed to line-of-sight)
- Shelter type must be on agency list of owned shelter types

#### **7.3.4      *Completeness***

Completeness is a measure of the presence and absence of features, their attributes and relationships. Completeness information includes feature omissions, selection criteria for inclusion of map features, generalization, definitions used and other rules used to derive the dataset.

For example, the completeness of the street names and address ranges on a transportation network can be tested by geocoding large, high quality (current and complete) data files against it, such as the tax assessor's data file of tax parcels.

#### **7.3.5      *Lineage***

A lineage report for a dataset provides information about the events, parameters and source data that were used to construct the dataset. The lineage report can also include information on source data publication dates, processes and responsible party contacts, purpose of the original data collection, and scale of the original map used for digitizing. Finally, the lineage report includes update activities, update dates and processing steps that have transformed the data. From time to time, different parts of the base map may be developed by different vendors or agencies. As a result, data within a base map may be derived from different applications or may have been added at different times. Data quality is highly impacted by the data source, by the procedures used to collect and derive data (lineage) and by when the data was updated (currentness). This issue may be particularly important between base map updates if scheduling software handled interim updates to route extensions or detours. The data that are transferred from the scheduling application may not match the base map and associated feature data to which it refers without the associated lineage and temporal attributes.

For the bus stop example above, the lineage information may include:

- Procedures for collecting individual fields of the attribute (e.g., assignment of a unique number for the bus stop ID, equipment and procedures to measure offset from corner curb and curb setback, etc.)
- Person(s) who collected and/or validated survey
- Update dates: survey update, inserted into database
- Associated transportation network (street address)

### 7.3.6 Data Quality Elements and Subcomponents

Table 7-2 summarizes the data quality elements along with their data quality sub-elements per International Standards Organization (ISO) 19113 (an international metadata standard) and the FGDC Content Standard for Digital Geospatial Metadata.

**Table 7-2: Elements and Categories of Data Quality (ISO 19113, FGDC-STD-001-1998)**

Element or Category	Sub-element
<b>Positional Accuracy</b>	Absolute or external accuracy—closeness of reported coordinate values to values accepted as or being true.
	Relative or internal accuracy—closeness of the relative positions of features in a dataset to their respective relative positions accepted as or as being true
	Gridded data position accuracy—closeness of gridded data position values to values accepted as or being true.
<b>Temporal Attribute Accuracy</b>	Accuracy of a time measurement—correctness of the temporal references of an item (reporting of error in time measurement).
	Temporal consistency—correctness of ordered events or sequences, if reported.
	Temporal validity—validity of data with respect to time.
<b>Thematic Attribute Accuracy</b>	Classification correctness—comparison of the classes assigned to features or their attributes to a universe of discourse (e.g., ground truth or reference dataset).
<b>Logical Consistency</b>	Conceptual consistency—adherence to rules of the conceptual schema.
	Domain consistency—adherence to rules of the conceptual schema.
	Format consistency—adherence of values to the value domains.
	Topological consistency—degree to which data are stored in accordance with the physical structure of the dataset.
	Qualitative attribute correctness—correctness of non-quantitative attributes.
	Quantitative attribute accuracy—accuracy of quantitative attributes.
<b>Completeness</b>	Commission—excess data present in a dataset.
	Omission—data absent from a dataset.

Element or Category	Sub-element
<b>Lineage</b>	Source information—list of sources, source map scale, short discussion of the information contributed by the source, time period for which the source dataset corresponds to the ground, and brief statement identifying the information contributed by the source of the dataset. Process Steps—information about each single event including the process description, the date and time when the event was completed, and contact information for the responsible party.

## 7.4 Recommended Practices

### 7.4.1 *Defining a Quality Process*

Although there is no silver bullet for implementing a quality process, there are key processes that if implemented consistently may improve the overall management of data within an organization.

Steps:

1. Describe your datasets, entities or data concepts in a data dictionary. Appendix F describes an approach to developing a data dictionary.
  - a. Unambiguously describe each field in each dataset
  - b. Unambiguously define the data format, coded-value domain and range domain for each field in each dataset.
  - c. Identify procedures for collecting, validating and storing of each dataset.
  - d. Identify requirements for updating each dataset and use the requirements to establish procedures for updating the datasets.
  - e. Provide enterprise-wide access to the data dictionary.
2. Develop “corporate” policies and standards for
  - procuring database management systems
  - procuring custom interfaces
  - developing internal applications on a limited set of software languages
  - application presentation format (i.e., corporate “look and feel”)
  - core data reference model
  - acceptable file formats (e.g., object graphics, vector graphics, word processing, geography)
  - Look up tables for street addressing and street aliases
  - Data entry instructions for the spatial data related to incident data (e.g., addresses)
3. Identify update requirements for applications that use the transportation network and transit feature data (e.g., Scheduling, Paratransit, Computer-aided Dispatch, Automated Passenger Counting system).
4. Employ data quality checks both when acquiring new data and when updating

The quality process may be institutionalized through key organization constructs such as:

- Documenting
  - Internal Data Policies and Standards
  - Enterprise IT and ITS standards
- Assigning Data Ownership
- Supporting a Data Committee and Organizational Data Surveys
- Documenting Metadata

The function of each of these institutional constructs is described in the sections below.

#### **7.4.2 Internal Data Policies and Standards**

Internal data standards are defined by an organization to describe and document the processes for managing the organization's core datasets. Important standards include the following:

- **Data Descriptions/Data Dictionary.** A data dictionary is a listing of the core data fields supported within your organization. For base map and transit features, it is assumed that each feature is composed of numerous attributes or fields.
  - Ensure consistent and unambiguous naming, identification, description and format requirements for each data/field (See Appendix F for more information on how to develop a data dictionary.)
- **Data Collection, Validation and Access Procedures.** Describe procedures for collecting, validating and accessing the dataset.
  - List the application/departments requirements that use the dataset
  - Validate using quality checks for logical consistency, attribute and positional accuracy, completeness, lineage and configuration management
- **Data Life Cycle Process.** Describe the life cycle and maintenance schedule for each dataset.
- **Data and File Interchange Formats.** Describe standard interchange formats or data sharing mechanism, including:
  - Association tables for specific applications/implementations
  - Document a Street Names and Alias table
  - File format standards (e.g., jpg vs. gif)
  - Transit feature data structures

#### **7.4.3 Enterprise IT and ITS Standards**

Similar to setting internal standards, industry standards should be used and enforced when procuring major systems if these standards apply to your organization's application requirements. Using industry standards can help minimize interoperability issues and the proliferation of computing environments and tool sets. The FTA's National ITS Architecture Policy on Transit Projects encourages transit agencies to adopt standards *where applicable*. There are an ever-growing number of standards for transit and spatial data. A list of these and how they apply to transit are discussed in Technical Memorandum #2, referenced in Section 7.5.

#### **7.4.4 Data Ownership**

In order for data quality policies to be enforced consistently, someone must be assigned responsibility for the dataset and its maintenance. An agency's information systems staff can take responsibility for data ownership or the responsibility can be delegated to a committee of appropriate stakeholders within the organization. In the latter case, the designated stakeholders are the primary data users, who can assess the accuracy, completeness, and fitness for use of the data. For example, the "ownership" of the Bus Stop Inventory may be assigned most appropriately to the organizational unit responsible for locating and managing the bus stops and their amenities. They would also need to work with other units (e.g., facilities maintenance, customer service, scheduling, Americans with Disabilities Act (ADA) compliance, Geographic Information System) to coordinate collection and management of the Bus Stop Inventory.

Clear requirements and procedures for data maintenance are essential. Distributing data ownership among stakeholders can be problematic if the stakeholders have a narrow view of requirements for data maintenance (i.e. their own). That is, if data "ownership" is assigned to positions deep into the organizational structure, those positions take on the responsibility to maintain an understanding of the requirements of all other users across the agency and to not put their own business needs ahead of other users.

#### **7.4.5 Data Committee and Organizational Data Surveys**

As a best practice, many organizations find that a data committee is an effective way of managing spatial data usage. Organizing a data committee is an institutional convention for gathering the major users and application developers that process the data to ensure that data collected meets their collective requirements. The committee typically develops an "official" enterprise description of each data, defines its update frequency and maintenance schedule, and identifies the resources needed to ensure its quality. Data processing requirements may be described, and redundancies and inconsistencies across the organization related to data *specification, production, delivery, use and update* may be eliminated.

A less dynamic means of acquiring similar information is through a data survey, where each organizational unit is asked to describe their use (including their applications' use) and requirements for core transit feature and transportation network data. This approach, unlike the committee approach, does not allow for interaction among the users. If used at all, it is best used in the preliminary stages of setting up a data committee.

#### **7.4.6 Metadata**

Effective creation and use of metadata is vital for successful location referencing efforts. Metadata are "data about data," that is, information that describes a data collection or part of a data collection including content, quality, condition/currency, organization, custodianship and distribution mechanism. The reasons for including metadata with a dataset are to:

- Organize and maintain the organization's investment
  - metadata are like maintenance records.
- Catalog datasets (in a standard way)

- metadata are like infrastructure management records.
- Provide information for data exchange
  - metadata are like interface specifications that include the purpose and “fitness for use” of the dataset.

Metadata contains the “self-description” of the spatial data file, including who developed the dataset, data quality, data accessibility, data descriptions, currency, lineage and more. The metadata may contain the data collection procedures, maintenance schedules, and statistical procedures.

Different types of metadata are needed for different audiences. Data users need a different level of documentation than do application developers. Users are interested in the quality related to using data. Developers need information about data relationships particularly with respect to integration points.

Advanced technologies tend to aid in the proliferation of operational data. The metadata serve as a way to index and catalog the datasets for integration, aggregation and processing. Metadata may be thought of as an institutional policy initiative. Defining a standard set of documentation and identifying where it is stored is a management issue; however, many software applications now read metadata to determine the most appropriate dataset for its use. Some of the data definition languages such as Extensible Markup Language (XML) can easily be adapted to embed metadata into the domain feature schemas. For example, a XML simple type that defines distance should include an attribute for units.

```
<xsd:simpleType name="SP-Distance" units="feet">
  <xsd:restriction base="xsd:unsignedInt">
    </xsd:restriction>
  </xsd:simpleType>
```

Because of its importance, metadata is treated in greater detail in Section 11. Section 11 focuses on the benefits, uses and best practices pertaining to metadata. Section 9 also includes a section that discusses self-describing interfaces that use the concept of “metadata” to transmit the interface data’s fitness for real-time use.

#### **7.4.7 Configuration Management**

Another important aspect of quality that is more of an overall data management issue than a specific data quality element is configuration management. Configuration management takes care of the information and processes pertaining to the update rates and data dependencies associated with maintaining individual or multiple, related datasets.

Example: An agency may maintain schedules for numerous “shake-ups” or “picks” simultaneously. These schedules are dependent on a set of routes, time points, patterns, stop points, fare zones, and other on-board data streams (e.g., annunciator event triggers). The procedures for managing these datasets depend on matching the metadata quality categories to the applications. These procedures are valid for applications, data stores, archives and warehouse systems, too. Establishing a configuration management

procedure for applications and application sharing is discussed in more detail in Section 9.3.

#### **7.4.8 Quality Checks**

Each of the data quality elements should be validated by Quality Checks. A quality checking process requires at least the definition of standardized quality measuring methods (metrics) and the assignment of quality levels to each measurement. The quality checking and quality concepts are relevant to the base map and its attributes (e.g., address ranges, one-way streets, topology, etc.) as well as transit feature data and supplemental datasets (e.g., demographic, orthophotos).

The quality checking procedure may be applied to a dataset as it passes through the phases of its life cycle. For example, ISO 19114 defines base map life cycle to include: specification, production, delivery, use and update. The transit feature life cycle may include creation or acquisition, validation, storage, integration, transformation, use, aggregation, reporting, display and update. Validation should be done before using the data in any way. First the collected information must be shown to be good; later there also may be steps to validate the collection process or the integration product.

#### **Transit Example: Bus Stop Data**

1. Collect (*acquire*) bus stop locations/order and core elements.
2. Download (*transform*) bus stop records into a management tool (e.g., database management system).
3. Validate and *store* data.
4. *Integrate (associate)* bus stop location with base map.
5. *Integrate (associate)* ordered sequence of stops to appropriate pattern(s).
6. *Select* data for other applications/analysis, such as a Customer Information System (CIS).
7. *Integrate* bus stop information into the CIS; match bus stop data with scheduling data (e.g., match patterns and trips that compose routes).
8. Reformat (*derive*) information for presentation to customers or Customer Service Representatives (CSR); *analyze* customer or CSR requests for nearest bus stop that includes a shelter to location.
9. *Report* and *display* bus stop information based on customer request.
10. Satisfied customer!

The steps in the transit feature life cycle can be grouped into higher-level phases: Steps 1-5 as Ingest, steps 6-7 as Propagate, step 8 as Abstract/Aggregate, and step 9 as Delivery. The following alternative steps 7-12 deal with phase of Update and Change.

7. Bus stop sequence was changed (so integration with route pattern is in error), and *report* to customer is inaccurate.
8. Error was detected by CSR and reported. *The error was based on selecting a location sensor that acquired data whose resolution was not sufficient to associate bus stops with route patterns.*
9. *An update* was sent to bus stop/route inventory management system.

10. Start over.
11. Bus stop locations *change* periodically.
12. Start from Step #1.

At each phase of this process, the quality of the data may change. Since the core data management system is the source for enterprise use, the quality checking process should be applied to the core dataset. (This assumption implies an enterprise approach to managing and distributing core data. Section 9.3 describes these practices in more detail.)

Assessing quality implies checking the quality of the data through tests, statistical sampling, and other methods. ISO 19114 describes two types of quality checking for each of the five quality elements. One method applies to format and structure, the other to semantic representation and interpretation.

1. *Format Quality Checking:* The extent to which the information contained is stored in the correct way.
2. *Semantic Quality Checking:* *The extent to which the information contained corresponds to reality. This cannot be checked before the format is error-free. (Is data meaning valid?)*

Format Quality Checking may be designed into your database management system by clearly defining the data structures as part of the data model, through referential integrity triggers or internal procedures. The process for checking format quality considers some of the aspects listed in ISO-GDF Table 7-3.

**Table 7-3. Error Groups and Examples of Format Quality Checking**

Error Group	Examples
Syntax errors	Use of characters not defined in a character set
	Incorrect record length
	Incorrect field length
	Field type violation
	Incorrect record padding
	Justification errors
Value errors	Incorrect record descriptors
	Incorrect record codes
	Incorrect feature codes
	Incorrect attribute codes
	Incorrect relationship codes
	Incorrect split identifier
	Value outside Min/Max range
Database integrity errors	Incorrect pointers
	Incorrect field counter values



Error Group	Examples
Topology errors	Level-0 objects sharing geometry
	Incorrect outline definition
	Disconnected line features
Value integrity errors	Incorrect feature—attribute relation
	Incorrect feature—relationship relation
	Incorrect relationship—attribute relation
	Coordinates outside section boundary

ISO-GDF as listed in Table 7-4 identifies some semantic checks and corresponding quality aspects.

**Table 7-4. Characteristics of Features That Can be Checked on Their Quality**

Check	Content of Check	Corresponding Quality Aspect
Positional Accuracy Check	Whether a Sample Feature has correct coordinates and shape	Positional accuracy
Topology Check on Level-1	<ul style="list-style-type: none"> <li>Whether a Level-1 Sample Feature is correctly present</li> <li>Whether a Level-1 Sample Feature has correct topological relations with other Sample Features</li> </ul>	Completeness and logical consistency
Topology Check on Level-2	Whether a Level-2 Sample Feature is composed of correct Level-1 and/or Level-2 features	Completeness and logical Consistency
Attribute Value Check	Whether a Sample Feature has correct attribute values	Attribute completeness and accuracy
Relationship Check	Whether a Sample Feature is involved in the correct relationships	Relationship Completeness and Correctness
Property Check	Whether a Sample Feature has the correct Properties	Associated attributes/relationship completeness

## 7.5 Resources

*Technical Memorandum #2: Issues related to Existing and Emerging Spatial Data Interoperability Standards*, Transit Standards Consortium, Location Referencing Guidebook Project, July 30, 2001. <http://www.tsd.org/tsc/lrg/index.htm> (see Final Report)

*Content Standards for Digital Geospatial Metadata Workbook*. Version 2.0; Federal Geographic Data Committee. US Geological Survey. Reston, VA, May 1, 2000  
[http://www.fgdc.gov/metadata/meta\\_workbook.html](http://www.fgdc.gov/metadata/meta_workbook.html)

*Content Standard for Digital Geospatial Metadata (CSDGM) Version 2 - FGDC-STD-001-1998*  
 Federal Geographic Data Committee. US Geological Survey. Reston, VA

<http://www.fgdc.gov/metadata/constan.html>

## **SECTION 8 ITS APPLICATIONS: POLICY REQUIREMENTS AND SPATIAL ISSUES**

### **8.1 Section Overview**

This section considers spatial data and location referencing issues from the perspective of implementing an Intelligent Transportation System (ITS) project. It provides information for transit managers, ITS project managers and technical staff to help ITS projects comply with the Federal Transit Administration's (FTA) National ITS Architecture Policy on Transit Projects, particularly the data sharing and systems engineering aspects.

#### **8.1.1 Problem**

It is an FTA policy goal, beneficial to transit, to facilitate intra-agency, multi-modal and regional data sharing. Many ITS projects and agency resources are not optimized or are put at risk because:

- Agency staff are unaware of the diversity of stakeholders and spatial data uses within transit and the interdependencies, hence they miss opportunities and requirements.
- The ITS project is not optimized from an agency-wide perspective, but is optimized to one, narrow business purpose; the application cannot be leveraged to provide additional value because project information cannot be integrated with the information and analysis tools developed for other areas.
- ITS project efforts often focus on application hardware/software and fail to address spatial data issues in timely manner and thorough manner; resulting in project delays, increased resource requirements and reduced functionality.
- Staff are unaware of the requirements of the FTA National ITS Architecture Policy on Transit Projects related to their ITS project, putting the project and funding at risk.
- Systems engineering processes are omitted or inconsistently utilized in a project.

#### **8.1.2 Purpose**

The purpose of this section is to provide information and resources to transit managers and ITS project teams to:

- Increase the likelihood of successfully implementing an ITS project that uses spatial data.
- Increase the value of the project to the organization, by potentially meeting more user needs and leveraging the resource investment.
- Help comply with the FTA National ITS Architecture Policy on Transit Projects (It should be noted that there are other required elements of the policy that do not pertain to spatial data and are not covered in this Guidebook.)

The tables in Appendix G, "Transit Applications/Functions that Use Spatial Data and Spatial Analysis Functions," are included to help transit meet the data sharing goals of FTA's policy and to help transit management and staff to:

- See a broad range of opportunities for using spatial data to improve transit.

- Identify other stakeholders, so agency-wide solutions may be developed.
- Leverage technology investments across transit by encouraging collaboration and data sharing.

The purpose of Section 8.2 is to help transit agencies benefit from systems engineering processes with respect to spatial data and to help comply with the FTA policy requirements for the use of systems engineering processes during ITS project implementations.

*Some transit applications that involve spatial data have special concerns. Section 8.3 addresses selected ITS issues, such as for a Bus Stop Inventory, on-board systems and Trip Itinerary Planning. Resources on systems engineering and a Bus Stop Inventory Guidebook may be found in Section 8.4.*

In addition to the information provided in this Section on ITS projects, ITS project teams will benefit from information provided in other sections, particularly Sections 5, 6 and 7 (Base Map, Transit Features and Quality).

## **8.2 Spatial Representation Issues Within a Systems Engineering Context**

This section will discuss, within the context of a systems engineering and project development life cycle approach, issues that relate to spatial data (e.g., base map and transit features) and spatial data sharing (e.g., for ITS applications).

**Systems engineering requirements and definition.** According to the FTA National ITS Architecture Policy on Transit Projects, “All ITS projects funded with Mass Transit Funds from the Highway Trust Fund shall be based on a systems engineering analysis.” Systems engineering is defined as “...an interdisciplinary collaborative approach to derive, evolve, and verify a life cycle balanced system solution that satisfies customer expectations and meets public acceptability.” Further, all ITS projects must include, at a minimum, the following systems engineering process elements [per <http://www.its.dot.gov/aconform/Arch2a.htm>]:

- Identification of applicable regional architecture portions
- Identification of agencies and roles (key stakeholders)
- Requirements analysis and specification
- Identification of relevant ITS standards and test procedures
- Alternatives analysis to meet requirements
- Procurement options
- System operations and management procedures, and resources
- Maintenance and update procedures and resources

Systems engineering processes help a project meet customer requirements, assess alternatives, lower risks and life cycle costs, and enhance the quality of the end result.

**Maintain an agency-wide and regional perspective.** The United States Department of Transportation (USDOT) guidance recommends that transit agency staff play a key role in the development of the systems engineering studies that are needed. The guidance is postulated on the premise that transit staff best understand the underlying requirements, institutional issues,

and needs of the “enterprise” and the region. The ITS project team should guard against making decisions that optimize (cost, schedule, design) for a single project, to the long-term detriment of overall operations and maintenance efforts.

In particular, the transportation network composed of the base map and transit features is an agency resource. Because the base map serves as foundational data and requires a significant investment in time and resources over its life cycle, spatial data management is an enterprise and/or regional issue and should not be determined by only one business area’s ITS project. For those reasons, any decision related to changing, migrating, duplicating, recompiling these core spatial datasets should be made to benefit the enterprise, with regional considerations as appropriate. The systems engineering process must assure this approach.

**Systems engineering practices and project life cycle phases.** The subsections below describe some issues and best practices related to spatial data within common phases in a systems development life cycle. Different organizations combine the various steps in a systems development life cycle in differing numbers of phases. At a minimum all these steps or phases, should be completed when implementing an ITS project, only the scale of the effort may vary, depending on the size of the project.

It is hoped that this section will provide a “spatial data” supplement to system engineering best practices. A systems engineering approach is a valuable tool for planning for ITS and spatial data deployments. Yet, much of the literature available is not specifically geared to transit, data management or particularly spatial data management. Many failed, delayed and overspent deployments occur because of “bad” data or insufficient data management procedures that are necessary to enable and implement application functionality and interfaces.

### **8.2.1      *Planning and Requirements Analysis Phases***

The planning and requirements analysis process includes a significant effort within the systems engineering approach. These phases focus on *What* is required. *How*, or implementation, is not part of this phase. *How* is the focus of the design phase. The Planning and Requirements Analysis phases include scoping and planning the project, a description of the stakeholders, requirements gathering and needs analysis, definition of the functional requirements, development of a Concept of Operations and mapping to the system architecture. During these phases, the operations activities should be generally specified so the necessary budget, tools and staff resources can be estimated for the operations and maintenance phase.

#### Stakeholder Involvement

The systems engineering approach encourages a collaborative environment, in which key stakeholders are involved in defining requirements, validating designs and testing functionality. With respect to spatial data, stakeholders may be identified as (1) providers of datasets that flow into the application and (2) users of data that are output from the application. These stakeholders include:

- Transit feature dataset suppliers
- Base map dataset supplier
- Application users (Appendix G highlights potential applications and users)

- Upstream data users
- Potential regional partners

Identification of existing data suppliers (transit features and base map) and datasets help define the “as-is” environment for the application.

### Concept of Operations

A Concept of Operations describes what the system should do from the user’s point of view in a concise, yet descriptive narrative. It includes how the user interacts with the system, and an overview of what is expected from the system. Appendix H includes a sample of a *Concept of Operations* document for an Itinerary Planning System that was developed as part of the Transit Advisory group of the Geospatial One-Stop program. The Concept of Operation’s focus was primarily on the spatial data requirements.

### Requirements Gathering and Functional Requirements Definition

Functional requirements list the user requirements for a system at a functional level. Appendix G includes a list of very high level functional requirements that apply to spatial data use and analysis within a wide range of transit business areas.

Table 8-1 highlights some key questions related to spatial data and the use of a proposed application that may be used to gather requirements. The questions can help develop functional requirements to support “fitness for use” of spatial data. For more complete information on defining requirements for the base map, please see Section 5, The Base Map. Section 6, Populating and Organizing the Base Map with Transit Features, provides additional information about transit features.

**Table 8-1. Key Requirements Questions on Spatial Data/Location Referencing (for proposed project)**

Functional Requirements Analysis	Key Questions
Base Map	<ul style="list-style-type: none"> <li>• What decisions within the application impose attribute or consistency requirements on the base map? <ul style="list-style-type: none"> <li>- Example: walking directions requires information on obstacles, barriers and access, as well as modal interconnects (e.g., walking and bicycle path connectivity to road network)</li> <li>- Example: driving directions and estimated driving times (for Paratransit) require topologically accurate road network, turn tables (allowable turns), speed limits and perhaps average driving time per time of day</li> </ul> </li> <li>• What type of analysis will be applied to the base map inside the application? (e.g., for distance measurement—line-of-sight or linear distance)</li> <li>• How current does the base map need to be?</li> </ul>

Functional Requirements Analysis	Key Questions
Transit Features	<ul style="list-style-type: none"> <li>• What transit features are required?</li> <li>• What location reference methods (LRM) are needed by the application and upstream data users? Do the LRMs need to be transformed between methods?</li> <li>• Will the feature set be mapped to more than one base map (e.g., another map in the region)?</li> <li>• If they are derived by field equipment (e.g., global positioning system (GPS)), do they include sufficient information to map them to a base map or transit feature?</li> <li>• For each feature, what attributes are needed to support the application (what decisions are made with the information)?</li> <li>• For each feature, what are the quality requirements?</li> <li>• Do some of the datasets depend on other datasets? (e.g., stop point amenities depend on stop points; integrated datasets will be dependent on the data that are integrated with them; complex spatial features such as a route may be dependent on other transit features. See Section 6 for more details.)</li> <li>• Do features need to be integrated together? How? (e.g., patterns are composed of an ordered sequence of time points, bus stops and event triggers)</li> <li>• Do they need to generate reports? What are the spatial characteristics that are needed in those reports? Can the data needed by the report be generated by the application?</li> <li>• With respect to performance measures, will they be associated with a transit feature? How? (e.g., passenger load is aggregated differently when associated with a trip, pattern and bus stop; running time is associated with a time point interval). [See Section 6 on Transit Features to see recommendations on how performance data may be associated with transit feature data.]</li> <li>• Are all the downstream (data providers) and upstream (data users) data interfaces supported with software modules to transport the data?</li> </ul>
Dynamic Position Information (e.g., Automated Vehicle Location (AVL) and field survey data)	<ul style="list-style-type: none"> <li>• If the location is derived from field equipment (e.g., GPS), how will it be used? What information will be needed to map to a base map or transit feature? (Will the location only need to be displayed, or will it be used for analysis and performance measures?)</li> <li>• What location accuracy is needed for upstream applications?</li> <li>• What data and diagnostics do you need to determine the quality of the data and the data collection device's health status?</li> </ul>

The questions above help identify the application's spatial data functional needs.

Most systems are not deployed in a vacuum; they work in an environment with other applications, pre-existing datasets and interfaces. As such, gathering requirements should also focus on the impacts of changing and leveraging these existing tools and datasets. Many of the questions in Table 8-1 are designed to uncover these relationships and effects.

### 8.2.2 *Alternative Analysis Phase*

An Alternatives Analysis assesses the feasibility of alternative design and implementation approaches and compares them to determine which alternative offers the best solution for the agency. The Alternatives Analysis helps an agency understand and balance trade-offs between

the functionality and benefits gained, the technical feasibility, ongoing maintenance and operations impacts, costs, and schedule requirements.

With respect to the transportation network (base map and transit features), the Alternatives Analysis assesses how well the spatial data meet the application requirements, weighed against the costs of meeting those requirements. For that reason, two major areas should be explored:

1. The “fitness for use” of the existing or proposed spatial datasets, including the base maps and transit feature data (See Table 8-2 for examples of some of the questions that should be answered in an Alternatives Analysis.)
2. The costs of the different approaches to acquiring the needed datasets (and any needed updates to the applications that use the data)

**Table 8-2. Alternative Analysis: Fitness for Use of Spatial Data**

Spatial Data	Key Questions to Ask to Assess Alternatives
Base Map	<ul style="list-style-type: none"> <li>• Can you use your existing base map? Can it be upgraded? Do you need to build or buy a base map (See base map evaluation criteria, a list of potential attributes and other base map discussion points in Section 5, Base Map). Examples of specific questions are included below</li> <li>• What is the lineage of the base map?</li> <li>• What attributes are supported? What are their formats?</li> <li>• Is the metadata complete and up-to-date?</li> <li>• How often is it updated? Is the base map changed (e.g., link and node identifiers) after an update process?</li> <li>• What file formats are supported?</li> <li>• Is special software supported to compile the base map for other applications and data users?</li> <li>• What datasets are associated with the current base map? What are the costs of associating them with another base map?</li> <li>• Must another base map be maintained for an interim period?</li> <li>• Will there ever be a need to exchange attribute or transit feature data between the multiple base maps?</li> <li>• What resources will be needed to ensure consistency between the two base maps and the data that are derived or associated from the applications that use each base map?</li> <li>• Can the processes to ensure consistency be synchronized so inconsistent data are not disseminated?</li> <li>• Who will assume responsibility for its maintenance?</li> <li>• Are there Intellectual Property (IP) issues related to updating and sharing attributes and network features? (See Section 4, General Procurement Guidelines.)</li> </ul>
“Patches”	<ul style="list-style-type: none"> <li>• Does the application handle patches that are not supported by the base map supplier? How?</li> <li>• Is it more efficient and cost effective to maintain patches outside the application and import them or to support redundant patches (developed both inside an application and by a central management process)?</li> <li>• Who will assume responsibility for its maintenance?</li> </ul>



<b>Spatial Data</b>	<b>Key Questions to Ask to Assess Alternatives</b>
Transit Features	<ul style="list-style-type: none"> <li>• What transit features are supported?</li> <li>• Is the definition of the transit feature consistent with the one required by the application?</li> <li>• What location referencing methods do they support?</li> <li>• To what base map is the feature referenced? How is it referenced and how is it stored? What are the resources (time, materials) necessary to perform the referencing (e.g., matching, query, etc.)?</li> <li>• Is the indexing scheme and format consistent with the one required by the application and upstream systems?</li> <li>• What are the attributes that are supported? Do they meet the application functional requirements (completeness, positional accuracy, etc.)? Are additional field surveys required?</li> <li>• How easy is the transit feature update process?</li> <li>• Can the features be integrated to support the needs of the application? What are the resources necessary to integrate the datasets?</li> <li>• If a new dataset is needed, who will assume responsibility for its maintenance?</li> <li>• Can the applications that generate features export the datasets?</li> </ul>
Dynamic Position Information (e.g., AVL and field survey data)	<ul style="list-style-type: none"> <li>• How many on-board devices are needed to support the necessary ITS functions? Can the data from a single navigation suite support all the on-board systems?</li> <li>• Does the navigation/survey sensor need to be integrated with other navigation devices in order to achieve the coverage, accuracy and operational utility? What is the right combination of technologies?</li> <li>• Can a single field survey support the stop point and route survey? How do future updates get obtained? Can the application vendor accommodate pre-existing processes?</li> <li>• How good are the diagnostic and error checking reports and procedures? Can you determine the data collection device's measurement error (accuracy)?</li> <li>• Will you need to post-process the location to apply corrections?</li> </ul>

### **8.2.3 Procurement Phase**

Section 4, General Procurement Guidelines, describes many of the issues to consider as part of the procurement process. Some of the key issues to review in the procurement documents related to spatial data include:

- Are industry and agency standards and quality requirements articulated in the specifications?
- How does the maintenance/warranty contract deal with maintenance schedules for application configuration data loading and compiling particularly base map and transit feature data?
- Are the spatial feature indices changed during the process?
- Can the agency upload the spatial datasets themselves? (Is the file format for upload open?) If so, what is the level of effort/resources that are involved with the process?
- How are “patches” and interim updates handled by the application?
- Does the maintenance/warranty contract and/or IP clause permit the agency to export the necessary information to other applications and users?

#### **8.2.4      *Design Phase***

In the design phase, the Project Team will develop and review design documents and dialog box mock-ups. There are a number of spatial data concerns within the design that should be explored:

- Are the requirements included in the design (traceability of the design back to the concept of operations, functional requirements and Regional Architecture commitments)?
- Are there changes (improvements or degradation) to spatial data quality based on internal spatial analysis algorithms? For example, does an application algorithm degrade the positional accuracy by transforming the location reference system or limiting the data precision?<sup>1</sup>
- Are the location referencing methods that are needed by analysts of the spatial data or application users supported by the application?
- Does the design support Agency standards and procedures? For example,
  - Do vendors support export and import interfaces or application programming interfaces using the agency’s data dictionary or the Transit Communication Interface Profile (TCIP) standards?
  - Are agency developed alias and other association tables integrated into the application as needed?
- Are the planned procedures for compiling/updating the base map manageable by the organization?
- Can information from the system be quickly and flexibly extracted and distributed for review and analysis?

#### **8.2.5      *Deployment Phase***

The deployment phase is intended to include the implementation (sometimes called development), testing, documentation development and system acceptance phases. Key issues related to spatial data in the deployment phase include:

- Actively managing the acquisition or development of required spatial data, in addition to implementing the ITS application hardware and software
- Testing the application’s sensitivity to “dirty” location (e.g., spatial and address) data
- Verifying the quality of any new or transformed spatial data
- Engaging users in testing the effectiveness of the:
  - Procedures to successfully import or export spatial data (Are the location referencing methods “interoperable”?)
  - Procedures and resources required for updating the spatial datasets
- Developing and updating spatial metadata as needed
- Developing criteria for system completion and acceptance

One of the most common reasons for ITS project implementation delays is that datasets needed for the application are not ready in time. For example, many Automated Passenger Counting

---

<sup>1</sup> This example implies that the accuracy of the base map or transit feature may not be propagated to the application output datasets.

(APC) deployments are delayed because of the lack of quality bus stop inventory data. The spatial dataset acquisition/update process should be scheduled with the ITS application deployment schedule in mind. The planning phase of the systems engineering approach will identify dependent datasets such as the bus stop inventory. Planning and budgeting for the resources necessary to complete these dependent core datasets is a benefit of the systems engineering process.

### **8.2.6 Operations and Maintenance Phase**

Assuming that a good systems engineering process defined the project requirements, and the project was implemented with robust interfaces, then the data update process will be the most frequent area of concern for the operations and maintenance phase. Update issues pertaining to spatial data include:

- *Frequency*: How often does the spatial data need to be updated, beyond the regular schedule, for interim changes and special conditions?
- *Scope*: How extensive are changes (does the entire dataset need to be updated for a single change)?
- *Impact*: What will happen if the change is not made? Will the system be “taken down” for an update? How long does it take to update the dataset, etc.?
- *Synchronization*: How to sequence updates of the application with the update cycles of the various internal or external datasets (e.g., base map, bus stop inventory, schedule data, etc.)? Section 9 discusses synchronizing dataset maintenance cycles.
- *Diagnostic Procedures*: Are there tools and processes for testing quality of updates, or for detecting, diagnosing and correcting problems?
- *Retaining Partners*: If data maintenance efforts are shared because data originates in different business areas or organizations, an ongoing effort needs to be made to retain the partners, communicate responsibilities and ensure consistent quality procedures.

Some agencies, such as King County Metro update their datasets on a daily basis (overnight). The majority of agencies update their spatial datasets quarterly when schedules are changed. Major ITS systems often fail because the resources needed to update the spatial datasets required to operate the application are not included in the operations and maintenance budget. Many trip itinerary planning applications were never deployed or turned off after the first year because the bus stop dataset could not be uploaded or maintained past the initial upload. Agencies must be aware of the resources that are needed to operate and maintain the datasets to protect these major systems investments.

## **8.3 Selected ITS Application Issues and Comments**

### **8.3.1 Bus Stop Inventory**

An up-to-date Bus Stop Inventory is an essential component in meeting the spatial data and location referencing requirements of many ITS applications. Once developed, it can be shared and used by many transit applications and regional organizations (e.g., cities, Metropolitan Planning Organizations). The inventory contains the transit agency’s active and inactive bus

stops, including the attributes and location references for each stop. In addition to the information on bus stops that is presented in Section 6 of this Guidebook, the “Bus Stop Inventory Guidebook” is an invaluable resource. (See Section 8.4 for a full reference to the Guidebook title and where to order it).

***Key Recommendations (from the Bus Stop Inventory Guidebook)***

- Define the limits of the Bus Stop Inventory (BSI)
- Plan on a testing phase of a limited dataset
- Define data quality that meets user needs
- Define location referencing methods, linkages, requirements
- Develop and maintain metadata
- Prepare long term management plans
- Institutionalize the BSI process
- Plan for inventory evolution

***Core Elements***

Some of the core elements identified in the Bus Stop Inventory Guidebook are included below with a selection of comments that help reduce location-referencing problems.

<b>Inventory Element/Group</b>	<b>Inventory Data Elements</b>	<b>Description/Comment</b>
Bus Stop Identification		Unique integer for each bus stop that was ever used by the transit agency.
Location Coordinates		
	Latitude	Decimal degrees with a minimum of five decimal places (44.12345). Five decimal places provide an accuracy of 1.11 meters or 43 inches.
	Longitude	Decimal degrees with a minimum of five decimal places (114.12345); (longitudinal error varies by latitude)
	Altitude	[optional] Represent by 10 <sup>th</sup> of meters as xxx.x and specify whether it is from offset from street level or Mean Sea Level
Intersection Relative Coordinates		
	On-street	May represent street by segmenting <ul style="list-style-type: none"> <li>▪ pre-direction</li> <li>▪ prefix</li> <li>▪ street name</li> <li>▪ street type (suffix)</li> <li>▪ post direction</li> </ul> (see <i>Geographic Information Framework Data Content Standards For Transportation Networks: Transit “Annex D Address Extension to the Transit Model”</i> (draft))
	At-street	See comments under On-street

<b>Inventory Element/Group</b>	<b>Inventory Data Elements</b>	<b>Description/Comment</b>
	Distance from intersection	In meters to 10 <sup>th</sup> of meters
	Direction of routes	Use NTCIP 1403 SCH-RouteDirectionName (north, south, east, west, southwest, southeast, northwest, northeast, inbound, outbound, circular, destination, clockwise, counterclockwise, name)
	Placement relative to intersection	Use NTCIP 1405 SP-RelativeDirection (Nearside, Farside, mid-block, at, between, farside-mid-block, nearside-mid-block, opposite)
	Bus position/bay	Use unique integer to designate bus positions are bays

### ***Additional Bus Stop Related Location Referencing Issues and Comments***

Some of the location referencing issues pertaining to stops are listed below with associated comments or recommendations.

<b>Representation Issues</b>	<b>Comments/Recommendations</b>
Transfer clusters/hubs	Represent each location in the hub or cluster as an individual bus stop.
Overlapping jurisdictions and sharing bus stops	Include identification of who uses and who owns bus stop (they may be different).
Define bus stop location	<p>From Jim Kemp (NJ Transit)</p> <p>“One of the more serious data model issues is that of data definition. For example, there are at least five sets of geographic coordinates associated with each known boarding or alighting location. They’re all different, and they’re all correct:</p> <ul style="list-style-type: none"> <li>▪ [...] “surveyed” position of the stop</li> <li>▪ [...] “apparent” position, or the position at which you have to represent the stop if you want it to appear at the proper place on a given GIS base map</li> <li>▪ [...] “nominal” positions of any landmarks the stop may be associated with</li> <li>▪ [...] “reported” position at which a given vehicle actually stopped at that stop on a given day, and</li> <li>▪ [...] “calibrated” position, which represents an average of the many individual position reports received for this location from vehicles in the field”</li> </ul>
Key Linkage	Tie bus stop data to scheduling (patterns and trips)

### 8.3.2 On-Board Systems

Among the most common on-board systems that use location information are:

- Automated Vehicle Location
- Automated Passenger Counting
- Schedule Adherence
- Route Adherence
- Automated Signage
- Annunciation Systems
- Fare Collection
- Transit Signal Priority

AVL, as an unadorned system, merely collects current location and time information. Transit Cooperative Research Program (TCRP) Synthesis 24 defines AVL as a system that “collects, processes, and communicates [transit vehicle] location information with other applications that need accurate and timely [positional] data.”

#### **Key Recommendations**

- Integrate on-board equipment with a single source for vehicle location (combined GPS and/or other AVL sensors) to minimize maintenance costs and conflicting location coordinates.
- Relate revenue vehicle location sensor measurements (latitude/longitude) to position along a pattern, route, or relative to a transit spatial feature (e.g., bus stop, time point).
- Accurate bus stop level information is critical; field collection of bus stop information is needed.
- In areas where there are bus stops, the base map needs to be as accurate as the GPS collected stops, otherwise the stops will “float” on the display map without an obvious link or relationship to the underlying street network.
- Determine how event triggers, such as for annunciators, will be located and their relationship to the bus stops (may affect requirements for field survey work)—see Table 8-3 for different types of event triggers.
- Integrate bus stop information with baseline schedule information.
  - Integration of trip level data with patterns, including actual distances between bus stops along pattern (e.g., turning distances along pattern cannot be calculated from base map).
- Associate performance data with transit features such as bus stops, time point intervals or patterns. Ensure authoritative instances of these features are used enterprise-wide (including on-board the revenue vehicle).

**Table 8-3. Example of Trigger Location Descriptions [from Miami-Dade Transit]**

Type of Function	Trigger Location Description
Annunciator	Offset from bus stop along each pattern
AVL/Schedule Adherence	Time point area (+/- 30 meters)—from pull-out to pull-in (block) (revenue and nonrevenue)
APC/Schedule	Stop point zones and time point areas (+/- 30 meters)—from

Adherence	pull-out to pull-in (revenue and nonrevenue)
Fare Collection (zone)	Zones are associated with bus stops/stations, location and zone number of bus stop/station.
Transit Signal Priority (TSP)	Location of signal controllers that support TSP.
Rail/Mover Annunciator (on-board and at station)	Physical trigger in track circuits.

### ***Special Considerations for On-board Applications***

#### *What is Bus Stop Location?*

Procedures and rules for documenting bus stop location are critical. Location is collected and compared on a number of occasions; errors and inconsistencies introduced as part of any of these collection efforts will affect the accuracy of the data collected. Section 8.3 under “Additional Bus Stop Related Location Referencing Issues and Comments” includes a description of some of the subtleties associated with describing position, each of these positions are used for different purposes within the passenger counting application. (Imagine if you needed to collect, maintain and process this information for every application on-board your revenue vehicle!) The Location Table described in Section 6.4.6 helps associate multiple measurements.

#### *Missed Stops (or Time Points)/Stops at Unknown Locations*

Missed stops (or time points) and stops at unknown locations along a route may also cause problems. Direction (e.g., orientation of the revenue vehicle) and distance between transit feature points along a pattern help applications determine how far along they are relative to their path of travel. In some cases, for example turning distances, the only way to collect accurate linear distances is to drive a bus along the path.

#### *Vehicle AVL Sensor Data*

In some locations and under certain types of weather conditions, sensor data may become unreliable or be subject to increased error. Collecting sensor health and confidence parameters may help in post-processing the data to determine if the measurement was degraded.

#### *APC Level of Effort*

To realize the potential of APC data to help a transit agency, at least one full-time equivalent (FTE) is needed to run the program and complete tasks such as the following:

- Review the accuracy of the bus stop location data and the alignment of the stops with the underlying base map. Request data and map improvements as needed.
- Insure that APC equipped coaches are appropriately assigned to trips for appropriate data coverage when the fleet is not 100 percent APC-equipped.
- Data gathering, review of diagnostic reports, data cleaning and other quality assurance activities required to transform the raw APC data into a “cleaned” dataset. In addition, work usually needs to be done to strengthen the APC data linkages to the transit feature data so that the APC data can be aggregated and queried for analyses.
- Data aggregation and metadata updates.
- Data analysis, reporting and display.

### **8.3.3 Customer Information System (Trip Itinerary Planning)**

Customer information systems are particularly sensitive to location referencing problems and often promote customer complaints when information is incorrect. Given that customer information systems are one of the most often implemented ITS systems, a sample Trip Itinerary Planning Use Case is included in Appendix H. The Use Case has three major sections: a brief concept of operations, an overview of functional requirements and the data requirements related to spatial data.

The Trip Itinerary Planning Use Case provides an excellent starting point for transit agencies that need to document system requirements for a new project and comply with FTA's National ITS Architecture Policy on Transit Projects. In addition, the Use Case highlights key spatial data and location reference needs.

### **8.3.4 Real-time Customer Information**

Real-time customer information provides the customer with the actual performance of revenue service. The information may be provided through various types of customer displays and devices including interactive voice response, active bus stop/station signs, on-board announcements and signage, web based services to Personal Traveler Devices or Internet. The information may be presented in a number of forms and mixed with scheduled, planned and actual (real-time) information. Different types of information include:

1. Estimated Time of Arrival
2. Range of Delay (actual time of delay)
3. Time Left Last Stop
4. Countdown
5. Actual Location (displayed on a map)

#### ***Considerations for Real-time Location and Time Information***

##### ***Data Transfer Latency***

Data transfer latency from the source (e.g., vehicle AVL) to the customer device should be considered in employing the information types. For example, if a vehicle is polled for its location every three minutes, the data presented to the user may be up to six minutes later than "real-time." Transit vehicle AVL data may need additional time tags to account for the latency. These may include: time of recording current location, time of transmission, time of reception.

##### ***Predicting Delay or Time of Arrival***

Some systems may incorporate actual traffic flow parameters to estimate travel time from the last reporting time/location to a destination. If this is the case, generating transit travel times using the same base map road segment as that associated with the traffic management center flow measurements will minimize errors in the estimated travel times. Typically, algorithms used to predict time of arrival or delay are based on a continuously changing time series. Collection of additional travel times (e.g., velocity, maximum travel speed along a link, historic aggregation of travel times, etc.) may improve the prediction algorithm.



### *Vehicle Location Displays*

The display of a vehicle on a base map may hide errors of estimation. For example, in the display, the map scale may be 1:2,000 (e.g., 1 centimeter on the map is equivalent to 2,000 meters on the ground) and the bus symbol may be 1:200. When overlaid, the bus symbol is an order of magnitude greater than the map (a 30 meter bus will be displayed as 300 meters long). Maps may be misleading.

### **8.3.5      *Ridematching Software***

Ridematching software for helping form vanpools and carpools also relies on base map data. The software needs complete and accurate street and address data for geocoding origins and destinations to make appropriate matches. In addition, it needs an underlying base map on which to display the results. The customer map should be able to display key landmarks, park-and-ride lots, transit centers and routes to facilitate transfers between transportation modes and finding rendezvous points.

Ideally, the Ridematch software should allow transit staff to import agency updates of the base map and transit feature data such as routes and stops. It can be prohibitively time consuming and expensive for ridematch staff to independently make corrections to the base map, address ranges, commuter lots, bus stops and route information.

### **8.3.6      *Transit Signal Priority***

For Transit Signal Priority projects, transit needs to coordinate with the appropriate local jurisdictions with respect to location data for traffic signals. A data sharing agreement helps ensure that transit has the most current information on traffic signals, controllers, and sensors, including their location and attributes. When a new program is initiated, ideally the traffic signal data should be able to be represented as a GIS layer on the transit agency's base map, so the relationships between routes, bus stops, signals and sensors can be analyzed and mapped.

Different versions of equipment and software must be tracked since the differences can introduce performance irregularities or errors.

### **8.3.7      *Paratransit***

Paratransit software also requires high quality street network and address data to facilitate customer pick-ups and drop-offs. Like many other transit applications, Paratransit is most effective when staff also know where bus routes, bus stops, rail stations and other transit facilities are located. Again, ideally the software should be able to upload agency updates to the base map data and the transit feature data as needed. An update process should not "lose" all the interim corrections since the last major update.

If the paratransit vehicles have AVL capabilities, some consideration should be given to the possibility of coordinating AVL equipment on bus, rail and paratransit. Many of the data diagnostics, error reports, data analyses and display issues have similar characteristics.

## 8.4 Resources

### System Engineering References

April 2002 monograph "Building Quality Intelligent Transportation Systems Through Systems Engineering," by Mitretek's Paul Gonzales.

[http://www.itsdocs.fhwa.dot.gov/JPODOCS/REPTS\\_TE/13620.html](http://www.itsdocs.fhwa.dot.gov/JPODOCS/REPTS_TE/13620.html)

“Functional Requirement Analysis”

[http://www.itsdocs.fhwa.dot.gov/JPODOCS/REPTS\\_TE/13621.html](http://www.itsdocs.fhwa.dot.gov/JPODOCS/REPTS_TE/13621.html), or

[http://www.its.dot.gov/aconform/docs/systems\\_engineering.doc](http://www.its.dot.gov/aconform/docs/systems_engineering.doc)

(On-line and classroom training) Introduction to Systems Engineering

<http://www.nhi.fhwa.dot.gov>

### Bus Stop Inventory Guidebook

*Bus Stop Inventory Best Practices and Recommended Procedures*. Bus Stop Inventory Task Force, Transit Standards Consortium, 19 September 2001.

Available from:

Transit Standards Consortium

91 West Clinton Avenue, Suite 2

Tenafly, New Jersey 07670

<http://www.tsconsortium.org>

### Transit Signal Priority

TSP Overview Sept 02, ITS America, publication 326983

[http://www.itsa.org/resources.nsf/Files/Transit\\_Signal\\_Priority/\\$file/Transit\\_Signal\\_Priority.pdf](http://www.itsa.org/resources.nsf/Files/Transit_Signal_Priority/$file/Transit_Signal_Priority.pdf)

### Travel Information: Guidance Documents published by FTA

FTA publications are available from FTA Transitweb <http://www.transitweb.its.dot.gov>

Customer Preferences for Transit ATIS: Research Report. (FTA-OH-26-7015-2003.1)

Guidance for Developing and Deploying Real-Time Traveler Information Systems for Transit.

## SECTION 9 SHARING BETWEEN APPLICATIONS: COMMON INTERFACES AND MAINTENANCE SYNCHRONIZATION

### 9.1 Overview

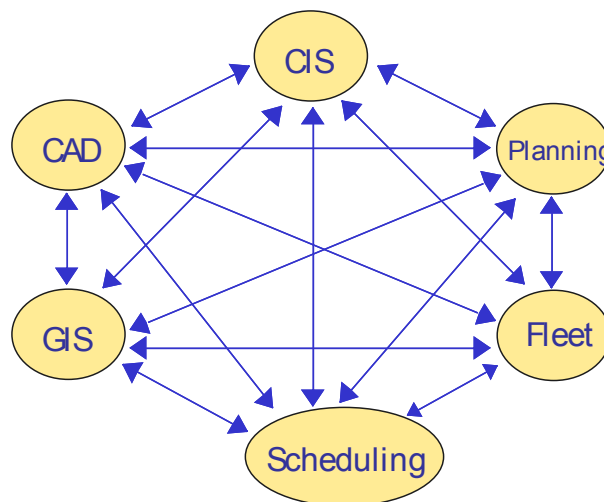
After a base map and transit feature data are available to a transit agency, there is an inevitable need to distribute and share the data within the organization and among its automated systems. For example, trip planning systems need base map, route and schedule data to operate. Similarly, planners need to combine automated passenger counting (APC) data with the base map and route alignments to assess ridership. Much of the data sharing and combining occurs through the use of data interfaces between applications.

This section identifies some of the key issues and best practices related to data interfaces. It discusses why transit project managers and technical staff need to understand and manage the synchronization of the dataset versions that are interfaced.

#### 9.1.1 Problem

The key objectives of this section are to describe best practices with respect to the unambiguous transfer of information between systems and to help transit staff ensure that the data throughout the organization are consistent and complete. There will always be a need to share spatial information, update base maps and attribute information, update core transit feature data, integrate transit feature data with geospatial features for analysis, and transfer real-time data for application transactions. Yet, many projects are delayed because the interfaces between applications are problematic. The interfaces do not work properly or the interfaces cannot be maintained over the long run. This world-view may be thought of as the “spaghetti model of interfaces” (see Figure 9-1) or more technically referred to as an application-centric approach.

**Figure 9-1 Application Centric Approach**



Even more problematic is ensuring *unambiguous* transfer of location information. A location reference managed in one application or data source may not be completely understood by another application. A bus stop may appear on the wrong corner and side of the street, or a pickup location may be located a block away from the actual address. These types of problems will cause inefficiencies and delay deployment of systems and analyses that rely on integrated data.

The quality and validity of spatial analyses, maps and Intelligent Transportation System (ITS) data are also affected by the dataset version that is used of the base map, schedule data or other transit information. Transit agencies usually create many versions of their datasets with location-related data. Street networks and bus stop inventories are constantly being updated. Generally, past, current and future versions of the service footprint are retained. Further complicating version management and control is the advent of real-time data in transit. Agencies need to decide which interim-detour changes to the service footprint and the street network need to be recorded, when and how.

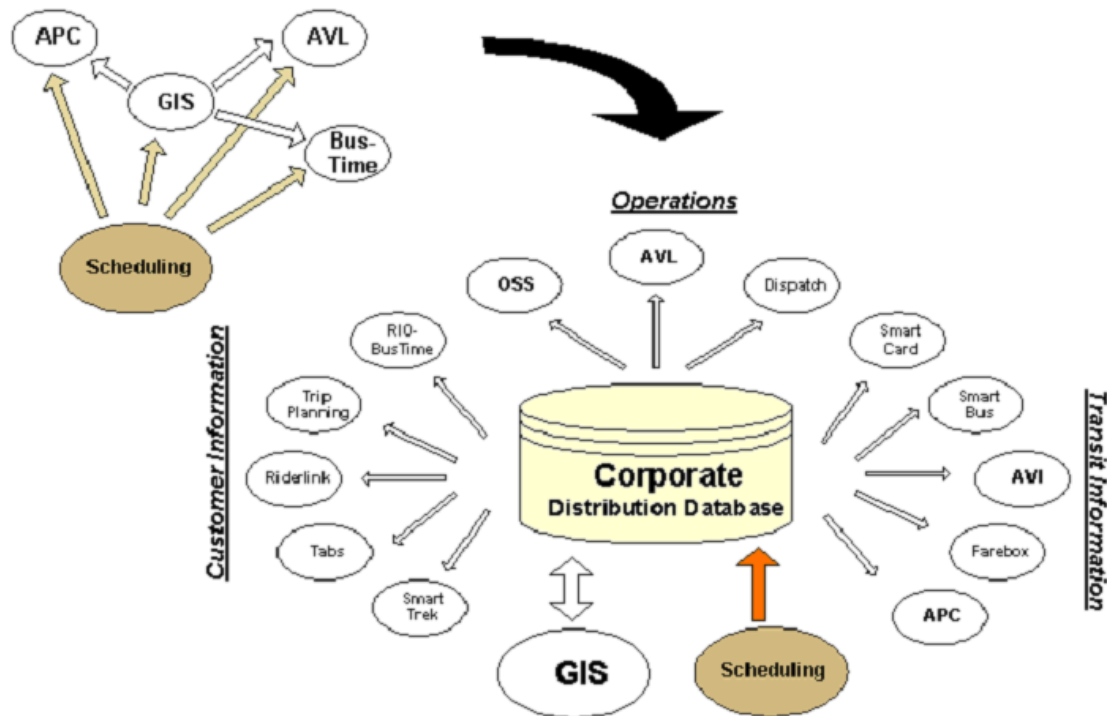
### **9.1.2 Purpose**

This section describes key enterprise management activities and technical procedures that support the implementation and maintenance of interfaces throughout the agency. There are four major parts of this section.

- Section 9.2 provides a summary of the management practices and other requirements that support successful interfaces and transformations of location referencing methods among applications and data sources.
- Section 9.3 describes the need for synchronizing core datasets so they are consistent in all systems that use them; this is accomplished by documenting and scheduling interface procedures.
- Section 9.4 describes a relatively new approach to sharing data that may be used by new systems and applications. The technologies, web services and middleware, are able to interface to existing systems and provide access to data for evolving applications.
- Section 9.5 describes a new technique that provides applications with information about interfaces, that is, the ability to include self-descriptive information or metadata along with the interface content.

This section makes the case for a data-centric or information enterprise approach to managing spatial data. The ***data-centric*** must replace the application-centric approach if we are to overcome “stove-pipe” implementation of applications. The section also discusses methods for preventing “stove-pipe” data by centralizing management of spatial data and location referencing procedures, and insulating the database while allowing for its evolution in response to changing business needs.

**Figure 9-2 Transit Corporate Data Evolution from Application Centric to Data Centric Approach** [Source: Michael Berman, King County Metro]



## 9.2 General Requirements That Support Interfaces and Location Referencing

The systems that share information through interfaces must support three levels of knowledge:

- Format description
  - Semantic description
  - Relationship of a “frame of reference” to other frames of reference
- These are the transformations necessary to convert from one version, representation or model to another

**Format description** refers to the organization of the “bits and bytes” and provides answers to questions such as the following: How is the feature represented? For example, is it represented an integer or character? How many bits (e.g., 16 bit) or characters are needed to represent it? For some applications, in what order are they transmitted (e.g., the Society of Automotive Engineers vehicle area network standard orders the bits differently than do personal computers)? Another important aspect of format description is the structure of information. How is an address defined? Do all your applications separate road number, pre-directional, street name, street type and post-direction into separate fields or are they combined? This too is considered part of the format description.

**Semantic description** refers to the meaning of the feature. Do different applications use the same definition? Does one application understand the definition the same way as another application? Is the definition of pattern that is being used the same in each application? For example, if one application or group is using as-driven distance and another is using calculated distance from a base map, then the two applications do not have the same understanding of the distance between stop points along a pattern. The metadata may help to distinguish differences in semantics when the same data element is used for multiple representations. When sharing data, metadata “touchpoints” described briefly in Section 9.5 and detailed in Section 11 support semantic description sharing.

**Relationship descriptions** refer to the relationship of the concept to the context or *frame of reference* from which it was created. A *frame of reference* may be based on a different representation of the earth’s surfaces, e.g., datum; it may be based on a linear versus a spherical representation of a network; or it may be dependent on a different time period in which the base map and features are valid. For example, one may view the timetable version for the spring service change as a different “frame of reference” than the winter schedule.

In using location referencing methods that are derived from heterogeneous sources (e.g., base maps, data collection sensors, orthophotos, etc.) and stored in various formats in vendor products, there are many ways to corrupt the relationship to a known frame of reference.

- The first and most often overlooked change in a frame of reference is the dataset version. Updates to a base map or bus stop inventory will alter the frame of reference. Even when you maintain a single base map and single source of core transit feature data as recommended in this Guidebook, ensuring that the same data is used by all end-user systems is a challenge. If the dataset is not consistent throughout the enterprise, then analysis using data derived from separate systems will be flawed; information provided to customers may be in error, and worse yet, studies presented to commissioners and government officials may cause embarrassment.
- Another frequently perplexing challenge is associating field survey or real-time data collection data to a single base map representation. For example, much of an agency’s operational data are collected using global positioning system (GPS) receivers, but the GPS points are not tied to the base map elements. On a map display, one may see a perfect overlay of the GPS collected data. Yet, as individual points without a base map foundation, there is no way to ascertain the direction, side of street, or corner to which the data refers. A time consuming process of geocoding the data must be performed to associate the GPS points to the base map, expending resources and delaying use of that data.

Many of the issues cited above are covered in previous sections. Certainly, the best practices provide techniques to deal with the format and semantic problems. This brief review of key issues is included here as a reminder that many factors need to be in place for an interface to operate successfully. In previous sections, relevant best practices include:

- Standardize on a single base map; store and manage the base map centrally

- Define a corporate transit feature dictionary and define an enterprise data model; store and manage the transit features centrally
- Establish a data quality process and enforce internal data standards/policies
- Assign responsibility for dataset development and maintenance
- Develop key lookup tables for location referencing methods (including street name alias table)

Describing data formats and definitions ensures semantic interoperability of spatial features. A single base map provides a standard frame of reference for integrating multiple data layers; a lookup table with equivalent location addresses for real-world features (such as transit time points and patterns) provides a mechanism for linking multiple base maps and application datasets that use different frames of references. Finally, a data quality process ensures that the data are consistent and complete, and the documentation describes its “fitness for use.”

### **Interface Types**

Since data are stored in relational database management systems (RDBMS), many applications and centralized data transactions occur through queries using Structured Query Language (SQL). These interfaces are dependent on the underlying data model. They may be built on the server side, called *stored procedures*, or integrated within an application (i.e., client-side), called *embedded procedures* or operations. The stored procedures are part of the database management system, and as such are managed centrally, used by many applications/clients, and mapped to the internal structure of the database; (changes to the database model may trigger a message to update a stored procedure). Middleware are neither stored nor embedded procedures, although they are more like stored procedures in that they are centrally managed and standardized for use by many applications/clients. (Middleware is discussed in Section 9.4.3 below.)

Generally, most data sharing occurs through a batch file transfer. A file is generated from one data source and retrieved, inserted and processed by an application. As transit agencies deploy more real-time decision support and customer information services, the need for interfaces that can support real-time transaction and control is increased. These applications will need faster data access technologies such as stored procedures and middleware. Section 9.4.3 discusses emerging tools that may make the distinction between batch file transfers and real-time transactions moot.

## **9.3 Synchronizing Dataset Maintenance Cycles**

The timing requirements of the agency’s business processes have an enormous impact on the schedule for updating software and maintaining datasets. Also, the coordination of update cycles may be further complicated by the need to align the efforts of three stakeholder groups. The three stakeholder groups that impact the creation of combined datasets are the:

- Vendor or individuals responsible for updating the base map
- Systems analysts or vendors responsible for updating the transit applications
- Agency business users

Internal service changes, base map upgrades, and application maintenance schedules may need to be synchronized so that a change in one set of data will not invalidate another or inaccurately display a transit feature location on a base map.

At least four types of spatial data changes must be considered when planning a data maintenance/update schedule to minimize the inconsistencies among applications and meet varying business needs.

1. Changes to transit patterns, time points, bus stops and routes.
2. Physical changes to the roadway network within the service area such as new road construction, road realignment, and changes in traffic patterns.
3. Changes to business data referenced to the base map such as the construction or removal of landmarks (apartment complexes, shopping malls, hospitals, etc.).
4. Changes to the base map, but not the real world, such as improved cartography.

The base map and transit feature data are enterprise-wide resources. The transit agency must come up with a plan and schedule for maintaining them. A Gantt Chart describing the update cycles and data dependencies can support the development of a comprehensive schedule for meeting base map, application and agency needs. The dependency constraints for the scheduling problem are as follows:

1. Update cycle for application software (fix bugs, new features). New versions of software may require different data formats or new data, or they generalize the data in a new way.
2. Schedule for getting updates of the base map from commercial map vendors or regional/internal staff who maintain the map.
3. Schedule for getting or generating recompiled versions of the base map for certain applications such as Scheduling, Trip Planning, Paratransit, and Computer Aided Dispatch.
4. The update dependencies of each application. This includes the dependency of one application on output data from another application and the dependency of each application on each of the four types of spatial data changes listed above.

Generally, when developing an application maintenance schedule, key questions that should be asked include:

1. What transit features and attributes are needed? What types of location referencing methods are needed (address, latitude/longitude, relationship to other features/attributes)?
2. How often are the major changes to service data made?
3. How often are interim changes made?
  - Do dependant applications need to be updated on the major change cycle only? Do they need the interim changes?
  - Do you have authority to make changes to transit feature locations yourself in the application (e.g., field survey collected latitude and longitude for bus stops and landmarks)?
4. How often is the base map changed?
  - Do you have authority to make changes in base map (e.g., add a road through a mall parking lot)?
5. How are applications updated?
  - a. Who updates base map



b. Who updates transit feature data (e.g., pattern alignment, bus stop)

The agency-wide benefits derived from coordinating schedules for updating the base map, transit features and applications are consistency across applications and avoidance of redundant data maintenance. When new data are added to the base map, then every application gets the same data. If new data are added to the base map to support new features of one or more application, then every application is loaded with the same data. New features must be sent to upstream applications (Computer Aided Dispatch, customer information) so that all applications use the same information. Major transit service changes will impact application products.

Some application products will be impacted by interim changes. For example, on-board announcements will be impacted immediately when a bus stop or pattern is moved. The announcements, because of the public exposure, will most likely need to be changed immediately. Schedule adherence collection will be affected by adjustments to a trip. The trip adjustment may wait until the next period if the data is post-processed, or a flag is associated with the data to show that it is “flawed.”

The update cycles for applications may depend on policy and public relations rationale. They also depend on the internal technical skills and workload, or vendor maintenance contracts and licenses.

### **9.3.1      *Base Map Update Process***

Base map updates ripple through applications that rely on or are geocoded to the base map features. Section 5 on Base Maps addressed many of the questions that should be asked as part of the base map acquisition process. The update process should carefully consider the needs of each application that uses the base map or data derived from the base map. Critical issues such as configuration management, special feature representation and interim changes are discussed in more detail below.

#### **Configuration Management**

Applications across the organization should use the same version of the same base map to ensure data consistency throughout the organization. Sometimes this is not possible, although, systems that must communicate with each other should have the same update cycle. For example, if the location of a bus stop moves, the following staff or systems should use the same updated information:

- Customer Service Representatives (CSRs) who provide bus stop location information, including the Itinerary Planning system, if any
- Interactive Voice Response (IVR) systems that provide bus stop related information to customers
- Next stop announcement systems that provide information over a revenue vehicle’s
- Agency or regional web sites that provide bus stop information
- Bus Stop Inventory (e.g., field location of bus stop sign)
- Automated Passenger Counting system that collects ridership information at the bus stop

Consistency in the data used by all applications means that everyone is using the same data across the organization. Consistency ensures that performance data collected and used, and information provided to the customer are of similar quality and may be compared while minimizing errors.

Yet, ensuring consistency may not achieve consistent accuracy. Some applications “recompile” the base map and its attributes to improve their internal performance. In so doing, they may change link and node identifiers or they may reduce data precision by aggregating, generalizing or imposing naming constraints. Thus, errors may still be introduced that degrade accuracy, even though the same data were originally passed to all the systems.

### **Special Base Map Features**

Special features may need to be inserted into the base map to support general and specific transit business needs. For example, transit routes tend to traverse mall parking lots and industrial parks that may not appear on commercial or public agency generated map products. As described in Section 5, there are ways of building “patches” for the base map. These patches may not be compiled with the commercial product. Many transit agencies use internal mapping tools from their application vendor to manually draw private streets into the application base map—*not a desirable solution*. As discussed in Section 5, this Guidebook recommends that the transit agency work with their base map supplier to include these private roads into the delivered product.

### **Interim Updates**

The timing of base map updates may not be frequent enough for the transit agency, for example, construction on a bridge that disrupts significant routes. Map vendors adopt different policies for when interim changes may be included in a map. If construction will last less than six months, then the vendor may not change capacity or flow. Yet, for transit, six months consists of a complete “shake up” or timetable version.

A work-around that violates core quality processes, yet undertaken by many organizations is to manually draw streets or change street status using the mapping tools supplied in their vendor application. This line work by definition is unique to the application. Caution should be taken when using application mapping tools; few transit applications support a base map export utility, or incorporate specialized functions to verify and validate the quality of the cartographic work performed.

**Best Practice:** Most agencies that adopt a policy where changes may be made in an application also adopt procedures for tracking the changes, for transferring those changes to a “permanent” version of the base map, and for re-associating transit features to that permanent version.

### **Questions to Ask About Base Map Loading and Updating**

Keeping in mind the issues on configuration management, special base map features, and interim changes there are additional questions that may be asked. Moreover, applications that “recompile” the base map may change the relationships and quality of the base map. The recompilation process and resulting format are exclusive to the application program and may be proprietary. As a result, known and unknown changes to the corporate base map may creep in as

a result of the internal workings of transit applications. So even if there are metadata on core datasets and the base map, if data is retrieved from an application, that data should include additional quality related metadata. Listed below are additional tests and questions to ask with respect to base map recompilation process.

- Who has responsibility for changing the base map? Is there an open loading/updating format?
- Can interim updates be made, or does the whole base map need to be updated at once?
- Do applications change the base feature parameters when compiled internally? That is, are topological relationships maintained when applications compile the base map? Is coordinate precision preserved? Are base feature identifiers changed? Are naming conventions preserved?
- What attributes are needed by the base map to support the application?
- Do the spatial algorithms preserve accuracy and precision, topological relationships? What errors are introduced as part of the analysis process?

The answers to these questions may require special treatment for converting the base map to select applications.

### **9.3.2      *Building a Base Map Data Configuration Management Plan***

A base map configuration management plan supports maintenance and update of a base map and feature data for dependent applications. The management plan identifies how to handle major and interim changes. It describes the transformations executed on the source data and the changes necessary to load the data into the respective applications. An example of the rules for transforming Tri-Met's base map data into the respective applications formats is included in Appendix I.

As an example, the Tri-Met configuration management plan (see Appendix I) consists of the following components:

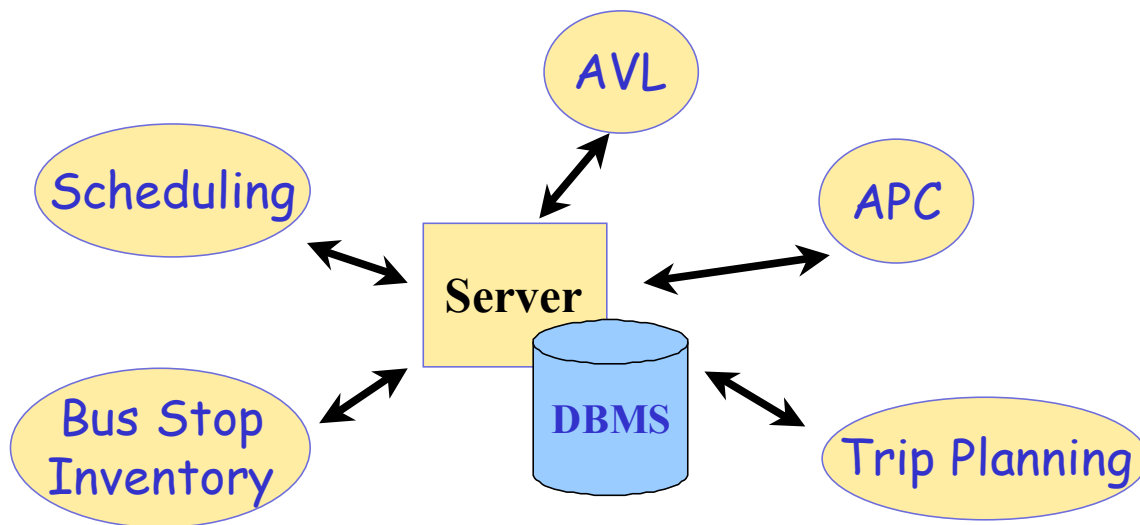
- Identification of dependent applications/users for each dataset
  - Includes the format for each application/user
- For each dataset, identification of dates and persons responsible for major and interim changes
- Notification of Change due to interim changes for dependant applications and users
- Tracking of changes to datasets (changes, date changed and person responsible)
- Transformation rules applied to data formats for dependant application and users
  - Includes names and locations of files, modules and source code that automatically make those changes

Although the Tri-Met Plan does not do so, a configuration management plan may be documented as a database where the time stamps, person responsible, and other metadata for each application is implemented in its own table.

## 9.4 Impact of Client/Server Models

### 9.4.1 Client-Server Approach and Issues

A typical computing architecture supported by many transit agencies is the Client-Server model, with a centralized database management system. Often, SQL is used to access data. Figure 9-3 illustrates this centralized data management approach. This approach is consistent with the data-centric model described and illustrated in Figure 9-2 above.

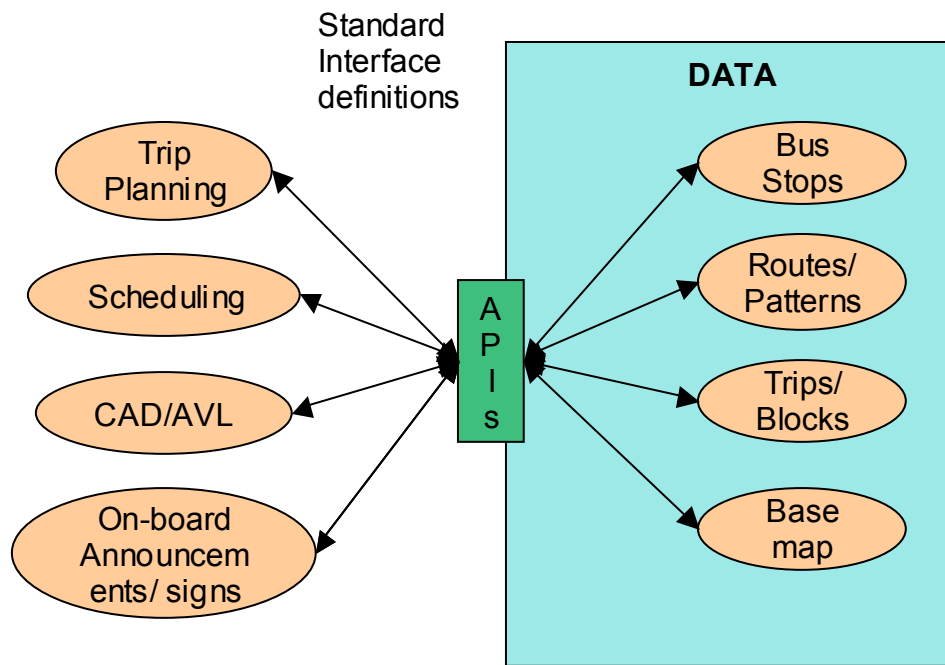


**Figure 9-3: Client/Server Approach to Support Application Needs for Core Data**

There are some drawbacks to this two-tier Client-Server approach. Although, many of the interfaces to each application may be reused, each application *embedded* SQL interface is unique. The interfaces are dependent on the data model remaining static. Although, mature databases are fairly stable, changes to the data model impact all the embedded SQL interfaces that use that relationship and related fields. Since applications require information in different formats, sequences and associations, the impact of a data model change may be significant if a key relationship is altered. Any change to the basic core model may also require an analysis of all the impacted application interfaces and the level of effort needed to make those changes.

### 9.4.2 Alternative Approach

An alternative approach to the centralized client/server approach is to standardize and centralize the interface definitions between the core database and clients that need information. This approach has the affect of isolating the data model from the interfaces, so changes to the data model do not “break” the application. Data access is managed through standard interface definitions (APIs) that are used by all applications as shown in Figure 9-4. The APIs may be implemented as stored procedures or as middleware.



**Figure 9-4: Standard Interface Approach to Support Application Needs for Core Data**

This approach may use a data repository or “portal” approach, or make use of application programming interfaces (APIs) to access the internal data. Extensible Markup Language (XML) standards support various methods for accessing and sharing information. In particular, an XML schema definition supports a domain vocabulary and inherent relations among data. Batch files or messages with data represented by the XML schema “tags” can be accessed, processed, transmitted or transformed using multiple methods. Most standard database management systems will automatically generate an XML content file based on a specified XML schema (such as Transit Communications Interface Profile, see Resources section).

The Geospatial One-Stop (GOS) effort (see References) provides a mechanism for outside systems to discover geospatial datasets. Although the transit model of GOS focused on three functional areas—regional customer information, regional bus stops, and rerouting service provision—any GOS compatible dataset may be accessed using various services such as:

- Catalog Service—search for data
- Web Map Service—geographic data rendered as image (maps)
- Web Feature Service (WFS)—storage and retrieval of geographic vector feature data (point/line/polygon)
- Web Coverage Service—storage and retrieval of gridded/continuous data
- Gazetteer Service—convert place names to coordinates

The WFS will provide the greatest benefit to transit applications. Applications built “on top of” the Web Map Service will be able to retrieve and display maps with transit features using this middleware module.

Other XML standards include transformation tools such as Extensible Stylesheet Language Transformations (XSLT) and messaging protocols such as Simple Object Access Protocol (SOAP). XSLT templates are designed to read a content file from one form and transform the content to another form (for display or file input), and SOAP serves as a messaging protocol supporting various ways of exchanging information including publish and subscribe, query processing, and more.

One method need not exclude use of other methods. The current computing architecture(s) supported by your organization is a critical factor in determining future enhancements and changes to the your enterprise data access methods. The methods used must also be driven by the business processes and meet the decision support requirements of each application. An IT strategic plan, as described in Section 3, will help provide key information on the type of framework for sharing spatial data that will meet the needs of your organization's business processes.

#### **9.4.3      *Middleware and Web Services***

As briefly introduced in Section 9.4.3, emerging information and transit industry standards will foster the development of middleware and web services.

The term *middleware* is used to describe software products that serve as the glue between two applications [<http://www.webopedia.com/TERM/M/middleware.html>]. Middleware is distinct from import and export features that may be built into applications. Two sides of middleware connect to applications and pass data between them. Middleware might be thought of as the engine that drives different vehicle chassis and uses different fuels and fluid products.

A typical transit middleware product may be a module that calculates estimated time of arrival. The software module may grab current bus location and route information, integrate it with real-time traffic conditions and provide the calculated ETA of the next bus to multiple presentation media such as an Interactive Voice Response System, Active Bus Stop Sign, fleet status map, trip planning software, kiosk and more.

Some middleware products link a database system to a Web server. This allows users to request data from the database using forms displayed on a Web browser, and enables the Web server to return dynamic Web pages based on the user's requests and profile. A Web Service is a type of Middleware that works using World Wide Web Consortium (W3C) standards. As defined by [*Web Services Architecture W3C Working Draft 14 November 2002* <http://www.w3.org/TR/ws-arch/>], "[a] Web service is a software system identified by a URI [Universal Resource Identifier], whose public interfaces and bindings are defined and described using XML. (The generic set of all names and addresses that refer to objects (typically on the Internet) and use short strings. The most common kinds of URI are URLs and relative URLs. URIs are defined in RFC 1630 [Definition adopted from W3 Specification]). Its definition can be discovered by other software systems. These systems may then interact with the Web service in a manner prescribed by its definition, using XML-based messages conveyed by Internet protocols." Implementation of web

services will allow transit agencies to post their data in a Transit Communication Interface Profile (TCIP) XML format and be retrieved by multiple web services such as regional 511 applications.

Standards will foster the implementation of middleware. Although, there are a few middleware products currently offered, some GIS vendors sell middleware tools to use for developing spatial analysis both for web applications (internet or intranet). As described above, the Geospatial One-Stop is developing an infrastructure to support deployment of middleware for various markets that use spatial data; transportation is one of those markets.

## **9.5 Application Integration Metadata**

Even with the assistance of middleware, one of the major obstacles to using data generated by other systems is understanding the quality of the data that are presented. Even with current methods of integrating applications, transit agencies struggle with finding the most successful ways to integrate new systems with new and legacy applications. Strategies for application integration aim to reduce the time, cost and complexity of building interactions among application systems (Schulte et al 2002d). These can be accomplished by using integration middleware tools, by documenting the application interfaces for future analysis of change impacts, and by reusing the interface definitions and messages, where possible.

**Best Practice:** Transit agencies should manage application integration using an enterprise-wide repository of integration metadata.

Integration metadata are data about the information that is communicated to and from sources and destinations outside of the application systems through “integration touchpoints” (Schulte et al. 2002b). Integration touchpoints may be thought of as guidelines or contracts for applications that use an operation or API. They are typically contracts on the use of an API between the sender and receiver systems. Integration metadata does not document the data model used within application systems, rather they describe the responsibilities and conditions that apply to the use of the interface.

Integration metadata contains information about the communication content, the identities of sender and receiver, and the interaction process mechanics and business implications. Table 9-1 contains a list of what might be included in the application integration metadata.

**Table 9-1. Useful Kinds of Application Integration Metadata (Schulte et al. 2002b)**

Category	Metadata	Description/Example
Communication Content	Data being sent	Schema of the transmitted objects.
	Syntax information	Tags, delimiters, field lengths, data types and representations, justification padding, truncation
	Validation rules	Such as for range checking, list of specific allowable values
	Transformation rules	Rules to enable mapping the transmitted object to a different schema.
	Housekeeping information	Message structure, message size, and multipart message formats
Identities	Where-from (source) Where-to (sink)	The specific application systems and their respective entry and exit points, files or databases.
	Contact information	Business unit or application project teams responsible for maintaining sources, sinks and their respective adapters.
Interaction process and business implications	Routing information	Rules for content-base routing, publishing, or subscriptions
	Communication model	Request/reply, one-way synchronous message, store-and-forward, and publish and subscribe.
	Technology for each source and sink	Transport mechanism and communication protocol
	Dialog characteristics	Application-level protocols that reflect behavior such as specifying sequence rules for acknowledgement, putting instructions into packets and handling errors.
	Quality of service requirements	Addressing latency, delivery guarantees, throughput, error rates and non-repudiation proof
	Throughput characteristics	Expected minimum or maximum volume per time period and frequency of distribution
	Message warehousing requirements	Retaining a copy of a message for a defined period of time for business activity monitoring, auditing and retransmission purposes; rules for determining which message to store and the duration of storage.
	Security requirements	Encryption, authentication, authorization, and related dialog characteristics

To build and manage integration metadata, transit agencies will need to establish a centralized application integration repository. The recommended strategy for developing the integration repository is to create metadata for all application data flows as they are created, modified or reused (Schulte 2002a, Schulte 2002d). Transit agencies should not commence to document the inter-application exchanges as they exist, but rather to develop the repository incrementally as each project is executed. The integration metadata may also be derived from the subsystems and data flows of an application/data architecture, for example integration metadata may be defined for the flows for which your agency is responsible in the regional ITS architecture.



A simple and crude repository is useful if it helps developers examine the impact of changes or to reuse metadata into applications. Having no documentation on interfaces or having scattered and incomplete information is worse than having a partial repository.

A number of middleware tools provide inter-application communication and integration metadata such as XML documents, message oriented middleware (MOM) messages; Component Object Model (COM), CORBA, and Java Remote Method Invocation (RMI) calls; remote procedure calls (RPCs) and web service Universal Descriptions, Discovery and Integration (UDDI) directory (Schulte 2002c, Schulte 2002b). For example, XML can easily be adapted to embed metadata into the domain feature schemas. An XML simple type that describes distance may include an attribute for units.

```
<xsd:simpleType name="SP-Distance" units="feet">
  <xsd:restriction base="xsd:unsignedInt">
    </xsd:restriction>
  </xsd:simpleType>
```

(Note: System performance and data size may increase by using a self-describing (metadata) format to identify units.)

## 9.6 Resources

### Integration Metadata References.

Schulte, R., M. Blecher, J. Thompson, W. Rishel, and M. Pezzini. (2002a). "Strategies for Managing Application Integration Metadata," Research Note T-15-5501. Gartner, Inc. March.

Schulte, R., W. Rishel, J. Thompson, and J. Sinur. (2002b). "Identifying Useful Kinds of Application Integration Metadata," Research Note TU-15-1931. Gartner, Inc. March.

Schulte, R. J. Thompson, M. Pezzini, W. Rishel, and M. Blechar. (2002c). "Tools and Sources for Application Integration Metadata," Research Note T-15-1932. Gartner, Inc. March.

Schulte, R., J. Thompson, W. Rishel, and B. Lheureux. (2002d). "Best Practices: Managing Application Integration Metadata," Research Note TG-15-1933. Gartner, Inc. March.

## **SECTION 10 EXTERNAL AND INTERNAL AGENCY COORDINATION**

### **10.1 Overview**

External and internal agency coordination minimizes location-referencing problems and facilitates cost effective spatial data sharing. By sharing the costs and resources to develop, maintain and distribute spatial data, transit agencies can frequently acquire more datasets and data elements. Further, the accuracy and completeness of the data are generally better. One of the intents of the recent Federal Transit Administration's (FTA) National ITS Architecture Policy on Transit Projects is to create agency coordination and data sharing.

A single base map or centerline file is the common framework for referencing transportation attributes both internally and externally. Other spatial and location data can be shared in addition to the base map, such as the transit routes and stops.

#### **10.1.1 Problem**

In the absence of internal and external coordination, redundant spatial data costs and data inconsistencies regularly occur. Location referencing problems are common. Time is wasted on tracking and resolving the data inconsistencies. Inefficient data integration and analysis efforts also consume resources unnecessarily.

To have successful external and internal agency coordination, a number of technical and political barriers such as the following may need to be overcome.

- The high cost of data ownership (data was paid for by one organization before data sharing was proposed) and the desire to recoup costs for data development efforts.
- The inability to arrive at a consensus regarding the valid values for attributes. Minimally this requires staff to become familiar with the data definitions of other agencies, or the agencies must develop common definitions.
- Transit applications that have embedded Geographic Information System (GIS) functions and/or base maps can pose significant problems if the transit agency does not specify easy, flexible methods for how data will be imported/exported from those applications.
- Tedious and costly import and export procedures can limit the frequency of updates. However, advancements in technology and in the industry are moving towards more integrated solutions which may support more frequent updates.
- Unavailability of easy to use technology to implement shared data maintenance. For some agencies, the data could not be displayed efficiently; and real-time data maintenance is still an issue.
- Maintenance cycles of shared data may not correspond between agencies. This may require agencies to perform off-cycle loads of data from other agencies, as the data become available.
- The ability to isolate and filter shared data when not needed. For example, mileage cost calculations would need to exclude the trips of transit routes operated by other properties.

- Shared data should be integrated with the base data. This could cause issues for systems that cannot use or even load data from other agencies.
- Management unwillingness to invest in spatial data infrastructure costs such as data maintenance, usually because they have not been presented with a cost-benefit analysis
- Stakeholder resistance to sharing improvements in data.
- Inability to put in place and consistently execute maintenance procedures.

### **10.1.2 Purpose**

This section discusses ways to partner with external agencies to share costs and gain better data. It also discusses ways of improving spatial data coordination internal to transit.

## **10.2 Coordinating with External Agencies**

### **10.2.1 External Data Sharing Approaches**

One way for transit agencies to cost-effectively obtain and maintain a base map is to partner with other agencies with complementary business needs for the same base map. Data sharing may occur among state and local governmental agencies, transportation agencies, Metropolitan Planning Organizations, emergency management organizations, and utilities within a region. Regional data sharing efforts are agency consortiums with formal partnerships among the agencies to leverage data resources. By way of partnering, regional transportation data can be shared at four possible levels (Ford and Widner 2000):

#### Level 1. Exchange of Digital Information.

- The most common form of data sharing involving the free exchange of datasets.
- Usually, no formal agreements between parties exist, other than possibly signing copyright or license agreements upon receipt of data.

#### Level 2. Written Agreement to Share Data and/or Work Together.

- Formalized agreement for cooperation between two or more agencies.
- Involves written agreements formalizing distribution policies, distribution format, native data format, usage, and output.
- Frequency and timeliness of updates are important considerations that must be formalized.
- Provides a mechanism for continuous data sharing through agency personnel changes.
- Establishes a foundation for further cooperation between the parties involved.

#### Level 3. Incorporation of Datasets by One Agency.

- One agency takes on the role of compiling data from all participating organizations. This agency may have responsibility over the geographic area comprising the participating organizations.
- The compiled data are returned as a finished product to the participating organizations.
- This level of regional data sharing requires that Levels 1 or 2 already exist among the participating parties.

- Often, the individual datasets from the participating organizations have differing data projections, data formats, accuracy levels, redundant features, and insufficient metadata. As a result, the quickest possible turnaround time for incorporating data is measured in days.

#### Level 4. Collaborative Updating of a Common Street Base.

- The most cooperative level of data sharing where participating agencies update information in a common source.
- The maintenance responsibilities are unique for each participant (non-overlapping) and comprehensive (no data without a responsible party).
- The participating agencies have an interest in at least some information provided by another participant.
- This approach provides a comprehensive dataset capable of supporting multiple applications at the participating agencies and yielding consistent results.
- Requires formal agreements on the data model, formats, road definitions, attribute sources, accuracy and data security.
- Incompatible software at participating agencies may be a limiting factor in data sharing.

The appropriate level of sharing may depend on the regional constraints and available funding sources.

### **10.2.2      *Transit Agency Examples of Regional Data Sharing***

In the San Francisco/Oakland area, over 20 transit agencies are sharing transit data to facilitate travel across modes and service providers in the region. In particular, data are shared to support regional trip planning.

Other examples of regional sharing of a base map occur at the Utah Transit Authority (UTA), in the three-county region around Portland, Oregon, King County (WA), Miami-Dade (FL) and Northern Virginia. In these areas, all county agencies have access to a single base map as a way to increase efficiency in communicating across agencies and to reduce redundancy in data collection and maintenance.

Utah Transit Authority (UTA) coordinates transit service in six counties and is the only transit agency in those counties. UTA began a regional commercial base map effort after they experienced problems sharing transit information with other public agencies that used different base maps. They also had an inability to share vendor data with local jurisdictions. The most difficulty arose for coordinating paratransit and trip planning because of the need to share attributes and geocoding.

In King County, county departments and cities are participating in the development of a regional transportation network and will share the maintenance responsibilities. The roadway network was developed by King County Road Services as photogrammetric line work to meet their specifications of 10-foot accuracy and consistency. Next, King County Metro Transit conflated attributes such as street names, addresses, functional class, and high occupancy vehicle (HOV) indicator from the previous base map to the new line work. Additional attributes such as travel

direction, turn restrictions, and speed limits are added from GDT's commercial product through a partnership with GDT. Finally, edges were added for commercial areas and additional modes such as rail, pedestrian walkways, bike paths, and ferry crossings.

The county is responsible for the unincorporated areas and the cities are responsible for the incorporated areas. Some agencies such as King County Metro Transit, Sound Transit, and the regional E911, crosscut incorporated and unincorporated areas. In addition, Metro Transit maintains the transportation network edges on Metro-owned property and is taking on the responsibility of maintaining the areas for non-participating cities.

The Virginia Department of Transportation (VDOT) and the Fairfax County Government have entered into an agreement to share and maintain street centerline information that will serve the needs of both agencies. The agreement relies on collaborative maintenance of the underlying linear datum. Based on the NCHRP 20-27 (2) road data model, the linear datum enables multiple cartographic representations and multiple linear referencing systems that service different application needs.

### **10.2.3      *Motivations for Regional Data Sharing Efforts***

There are several reasons why a group of agencies in a region may get together to share data.

- Political mandate. For example, the County Executive of King County, Washington ordered all county agencies to work together to develop a transportation network that satisfies all county agency needs.
- Defensible regional decisions. Regional partnerships may stem from the realization that the regional decisions are based on information provided by the subregions. Cities, towns, and villages are the most knowledgeable about their areas and have the most to gain from maintaining their pieces of the regional network.
- Policy compliance. Transit agencies must comply with the United States Department of Transportation (USDOT) policy to develop a regional ITS architecture and share data where possible.
- Reduced direct cost for data acquisition. Agencies in a consortium/partnership can pool funds to acquire (purchase or develop) base data that are common to regional agencies. For example, individual agencies in the Portland, Oregon area region pooled resources to purchase digital orthophotography for the entire three county area. Regional partners can also pool funds to acquire regional data from commercial data vendor. A commercial base map with regional coverage may provide a starting point for coordination among multiple agencies in the region.
- Shared/distributed responsibility for maintaining data. Participation in a consortium can reduce a transit agency's burden of developing and maintaining data. A regional partnership can provide an opportunity to get access to additional data such as the most recent new subdivision and streets, corrections from E911 staff that are in the field, and

new data layers such as rail lines, waterways and hydrography for assisting with potential pollution impacts.

- Access to additional datasets. Transit is usually unable to cost-justify developing some expensive datasets that are used less frequently, such as County parcel data. If shared with other agencies, the parcel data can be used when acquiring land to develop new facilities. Similarly, watershed data can be shared to evaluate possible pollution impacts.
- More powerful ad hoc reporting. A shared base map and accessible metadata decreases issues associated with ad-hoc data sharing and reporting (everyone in the region and state is utilizing the same consistent datasets including 911)
- More frequent updates. A regional consortium can work together to provide a dynamic data repository, accessible via Internet, for data maintenance, downloads and reporting.

Despite compelling reasons to share data and the availability of technology to assist the process, the success of data sharing efforts is highly dependent upon the commitment of the individual people involved. Data sharing occurs when people believe it is the right thing to do.

#### **10.2.4      *Important Data Partners for a Transit Agencies***

Transit agencies can benefit from participating in regional data sharing efforts. Regional data sharing efforts are particularly beneficial to transit agencies if certain other data partners are involved.

- E911 partners tend to have good quality data on geocoded addresses and high quality alias tables for street names. These can be valuable for Paratransit and Trip Planning systems. E911 feedback to sources increases accuracy for everyone (including alias tables for the centerline file). However, sometimes there are political and financial barriers for many public safety groups that use the same data files and still remain in sync with outside agencies. For example, E911 partners have demanding data needs such as near real-time updates of the transportation network. Their requirements need to come first. E911 might use a more current version of the base map, while transit may be able to function with less frequent updates. Data sharing with E911 partners can be highly beneficial but not without cost. Many E911 Computer Aided Dispatch (CAD) systems include complex business rules in the data models. Also some E911 data, especially addresses, may be proprietary.
- City and County departments of construction and land use (permitting of new subdivisions and streets).
- Metropolitan Planning Organizations (MPOs) usually have general responsibility for regional data. The MPOs support base map usefulness by frequently being the source of datasets on zip code areas, census tracts, political boundaries and county boundaries. An

MPO may help coordinate efforts to develop base maps, develop maintenance approaches and coordinate distribution of regional base map updates.

- Public utilities frequently invest heavily in maps to track their customer sites and facilities.

#### **10.2.5      *Impacts on Transit Agencies***

Data sharing arrangements may require new data processes to import data from and export data to the regional dataset. Participation in a regional data sharing effort may include additional costs for revamping internal business and data processes. The agency will need to reorganize its coordination of data so that all transit systems within the agency utilize the new regional data when it is available. Some agencies centralize the agency-wide update process to ensure consistency.

Because of inadequate standards in the transit industry, preparation of new base map datasets for internal transit systems is often time consuming and costly (i.e., conversions into different projections, file formats, character limitations, etc.). For example, when the transit agency receives a new centerline file and other data from its regional partners; the transit agency must update the internal databases for its various software applications, such as trip planning, paratransit, scheduling, Automated Vehicle Location (AVL), and the location reference of transit features. These costs and impacts can be reduced by phasing in data and procedural standards.

Some data processing is required to prepare transit datasets for delivery to the regionally-shared database. In the Portland area, because Tri-Met's transit data are referenced to the regional centerline file used by other agencies, the other agencies can receive and use a wide range of transit data. For example, Tri-Met delivers the following to its regional partners:

- Bus Routes are generated as linear events and delivered as shapepoints
- Bus Stops and Stop Amenities are stored in relational tables with linear coordinates, XY coordinates, and Unique ID that relates to amenity tables, historical records, and accidents
- Light Rail Alignment is maintained in the regional centerline file
- Transit Centers (TC) are maintained as polygons to aggregate stop level data at the TC level (i.e., costs)
- Park-and-Ride Lots, Light Rail Stops, Tri-Met Facilities, are maintained as points
- Service Areas and Fare Zones are maintained as polygons
- Schedule Information including the temporal data

Tri-Met has developed a data publication approach for sharing spatial data with other agencies that minimizes the number of shape files that must be created. Bus stops are published to the public in a shape file format as multiple points located in one place if more than one route uses a bus stop. For example, a transit agency may have 3,000 unique bus stops, but the published shape file might include 8,000 bus stop records. A bus stop's unique identifier may be listed in the table several times, each time with the stop's "x,y coordinate" and a different route identifier and route direction, to indicate all the routes that use the stop. This approach allows tabular data

to be sent and avoids the need to send multiple shape files for different analyses or complicated bus stop sequencing information by route pattern.

### **10.2.6      *Critical Success Factors***

The success of a regional data sharing partnership depends on its organizational and administrative structure. Critically essential are structures for coordinating and managing the partnership and its collective data.

- Each agency partner who is a data provider must have a business need for obtaining access to data that is being provided by another partner. This characteristic is essential for a functioning consortium of data maintainers. Partnerships may include partners who are exclusively data users. Obviously, the partnership cannot exist without participation from the data providers and maintainers.
- The consortium must have a designated lead agency to oversee development of partners, agreements, and the supporting information infrastructure. The consortium will need to determine the rules for data maintenance, approve changes to the database or application, and reconcile any cross-jurisdictional issues. For example, in King County (WA), Metro Transit took on the responsibility of managing the consortium.
- The consortium partners must agree on the allocation of responsibilities. Upper management of each participating agency must commit to supporting their agency's participation.
- The partners must agree on the content specifications of the entities in the shared dataset. The contents of the shared database must meet the business requirements of all the partners. Accuracy and completeness of each data layer or attribute must satisfy the most stringent user and most demanding application.
- The partners must agree on the logical design of the database, the location referencing methods, how data will be partitioned into layers, and how data will be related.
- Rules and well-defined processes are needed for data maintenance that insures the accuracy and completeness of the data. The partnership must establish formal agreements covering details such as who is responsible for maintaining what data, how frequently the updates will be made available to other partners and standard operating procedures.
- A commitment and a foundation for long-term maintenance are needed. The partnership cannot survive without commitment from the partners to provide ongoing and long-term maintenance of their allocation of responsibilities. Participating agencies will need to commit and plan for resource needs.
- Mechanisms for update and recompilation of the regional dataset must be designed and implemented. The shared dataset will need to be updated on an agreed-upon schedule



and as significant changes occur. The partnership will need to have a designated partner who is responsible for a centralized application. Portland Oregon's Regional Council of Governments (Metro) assumes the responsibility for integrating data into seamless regional mapping layers and distributing them on a quarterly basis. King County (WA) has a multi-user data maintenance application developed in ArcGIS with user-friendly interfaces for maintaining the transportation network attributes on the fly as needed. The centrally-stored transportation network will be available to the data maintenance partners.

- A mechanism for the convenient distribution of the regional dataset is very important. Somehow the shared dataset will need to be made available to the partners. The means for distribution needs to be identified and developed. In Portland, Tri-Met assumes the responsibility for distributing the updated regional database to the regional users on a quarterly basis via a CD. In King County (WA) the centrally-stored transportation network is expected to be available through a technology that will facilitate a two-directional exchange of deltas between the master database and the client copy of the database (termed replication). The deltas are exchanged using very small Extensible Markup Language (XML) files that do not require large bandwidth. Consortium partners can obtain an updated version as needed anytime after changes are made.

### **10.3 Coordinating Internally Within the Agency**

#### **10.3.1 *Spatial Data Sharing Internally Among Organizational Units***

Highly specialized transit applications designed for each of the major organizational units are commonly developed in a proprietary format for a "stand-alone" implementation and each has its own manner of dealing with location information. This results in redundant system maintenance efforts, data synchronization problems, increased opportunities for errors, and significantly increased Information Technology (IT) costs.

Currently, transit agencies are forced to expend resources synchronizing data stored in these separate systems or they must be satisfied with the limitations of a single software solution provided by one vendor.

It is difficult to streamline the sharing of information between the spatial data that is centrally maintained in a GIS, and those datasets developed and managed by the individual proprietary software modules.

Until a time of true data integration, transit agencies can minimize location-referencing problems and facilitate spatial data sharing among transit business areas and applications through several fundamental rules:

- Utilize a common base map throughout the agency to increase consistency, decrease redundancy, and reduce costs. See Section 5, Base Map, for additional detail.
- Define a single source for base map conversion and distribution within an agency.

- Have a clearly documented and understood methodology for updating all base maps within the agency to minimize and manage inconsistencies.
- Maintain a central core set of data that can be distributed and utilized by all systems and users through the agency.

The development of a coordinated IT/ITS Strategic Plan will also help resolve spatial data development, maintenance and sharing issues. Agency budget staff and procurement staff can help support an enterprise-wide approach to core data.

### **10.3.2 Stakeholders**

A few examples of internal transit organizational units that utilize spatial data in their business processes are included below. Appendix G includes a more comprehensive list of transit business area and their potential uses of spatial data.

1. Planning and Transportation for general planning and analysis of routes, stops and schedules
2. Operations to locate vehicles using AVL
3. Marketing to analyze Origin and Destination studies and locate businesses for Employer Pass Sales programs
4. Graphics to generate schedule maps, system maps, various geographic displays
5. Facilities Maintenance to track inventories or generate cleaning routes
6. Light Rail Operations
7. Security

Examples of transit systems or applications that may require spatial data, or a centerline file, for production:

1. Scheduling
2. Paratransit Reservations and Dispatch
3. Trip Planning
4. Dispatch/Automated Vehicle Locators (AVLs)
5. Automated Passenger Counters (APCs)
6. Ridematching
7. Ridership Modeling and Forecasting

### **10.3.3 Procedural and Technical Considerations**

#### Data Development

To meet the requirements for the systems and users within your agency, you must approach it from the standpoint of defining and meeting the most accurate data requirements that are needed. Cost-benefit analyses and agency-wide management concurrence help define needed datasets and their appropriate levels of accuracy. (Section 6 includes more detailed information on the development of transit feature data.)

1. Identifying stakeholders in agency
2. Identify stakeholder data owned so that others would want information to share

3. Develop one Bus Stop Inventory
4. Agree upon common definitions of core data elements
5. Develop an internal Data Dictionary

#### Data Maintenance

In the past, most transit applications required or assumed that one particular workgroup would be maintaining the transit dataset used within the application. This may not be the best approach, now that tools allow distributed updating of individual data elements within a dataset. The street data and other mapping attributes are most commonly maintained outside the agency and imported in as mapping layers. In addition, stops may be maintained outside of the Scheduling unit, in a group that urgently needs up-to-date bus stop information.

Finally, it should be recognized that spatial information is best maintained in a GIS.

#### Data Integration

It is important to facilitate linkages between databases to minimize complex data loading processes and the need for data synchronization.

#### Data Publication

Spatial information maintained in a GIS, and Schedule information, which provide the temporal aspect, must be processed together to produce a publishable itinerary dataset for use in information systems throughout the agency.

#### Data Analysis

It is important when performing analysis that departments throughout the agency are utilizing the same datasets and documenting their procedures to ensure that the numbers are consistent and reproducible. Also helpful are a common graphical user interface for data queries and analyses, and access to metadata that describes available data and its limitations.

The analysis of APC and AVL data require that the data are cleaned, aggregated, and populated in tables with proper documentation and linkages to the spatial objects for geographic analysis.

Because each transit application has its own way of dealing with location, it is often difficult to export data from one application to import into a GIS for analysis. This process is often tedious and time-consuming, and requires the process to be repeated for up-to-date analysis in the future. The use of a Location Table, as described in Section 6, facilitates the process.

### **10.3.4 Recommendations**

In summary, best practices for sharing spatial data to internal and external stakeholders include the following recommendations.

- Have a strong champion within the organization to promote coordination.
- Ensure that the champion has a global perspective and represent all business owners.
- Have an integrated IT/ITS Strategic Plan.
- Identify user needs; be cognizant of the need to go beyond one area's data needs.

- Use cost-benefit analyses and management consensus to determine data elements and accuracy levels. Until data are cheaper, a transit agency can not meet everyone's requirements (e.g., engineering accuracy, cadastral; unshared costs are typically too high for the number of uses).
- Identify stakeholders early and get their buy-in and needs *early*.
- Initiate a grassroots movement among each of the stakeholder groups; management will be more likely to accept costs and changing business practices when they hear from their line-workers.
- Agency coordination needs to extend into data maintenance; sharing in ownership and responsibility.
- Distribute maintenance throughout the agency to data stewards with a vested interest in the quality of the data.
- Develop a Responsibility Matrix to facilitate understanding of tasks and availability of staff.
- Assign responsibility for planning and implementing data flows and data integration.
- Have Geographic Information System professionals oversee the design and maintenance of the spatial data..
- Also, most systems are project-oriented; within a project the maintenance requirements are not cognizant to the data steward; data maintenance and organization required for supporting the data generated or required by the systems.

## 10.4 Resources

*Ford, B.J. and D.K. Widner. (2000). "Shared geography: Building a common street centerline resource to serve state and county governments." In Proceedings, URISA Annual Conference. Orlando, FL. August.*

## **SECTION 11 THE IMPORTANCE AND ROLE OF METADATA IN EFFECTIVE SPATIAL DATA USAGE**

### **11.1 Section Overview**

This section describes the importance of metadata to a transit organization. It provides a definition of metadata, highlights its benefits and identifies the federal mandate for metadata if federal funds are used to create spatial data. More technical information is also provided for analysts and other interested individuals on metadata standards, what to include in metadata records, priorities for implementing metadata for the first time, transit agency examples and metadata related resources.

#### **11.1.1 Definition of Metadata**

Metadata is a “self-description” about the content, quality, condition and other characteristics of a single piece of data or entire file of data. In other words, metadata is data about the data, such as its date of collection, format and accuracy. Metadata provides the necessary information for humans to understand the meaning and context of the data, as well as how it is expressed or represented. All these must be described in order to manage and use data effectively.

Metadata tells about the format, data quality, history and availability of a single piece of data or a data file. Metadata concepts have evolved for managing information packages such as documents, images, audio streams, videos, engineering drawings, and Geographic Information System (GIS) data files. By reading metadata, a person can find data containing the feature class they need and evaluate its suitability for use. The metadata answers question such as:

- What is the data format?
- Where did the data originate?
- When were the data last updated?
- What is the purpose of the data?
- What is the scale of the data?
- What steps were followed to create it?
- What attributes does it contain?
- How are the data projected?
- What geographic area does it cover?
- Are there any exceptions or special conditions within the set of data?
- How can the spatial data be obtained?
- Will it cost me anything?
- Who do I contact for more information?

Metadata can apply to a whole dataset, or just to one piece of data. For example, metadata such as cost, scale, format and contact information typically apply to a dataset, while metadata such as currency and attribution may apply to just one specific piece of data.

Standards for creating and maintaining metadata are available, and common GIS packages offer tools for metadata. This section describes the benefits of metadata, the available tools, and some basic guidelines for how to set up an effective metadata system.

### **11.1.2      *Problem***

In the absence of metadata, many customer service options and internal ad hoc reporting options are not possible. In addition, a transit agency without metadata is less flexible and efficient. It runs the risk of publishing erroneous information or making decisions based on bad information. Customers and staff need to know what information is available and how it can be used. The scope and quality of the metadata affect how quickly queries and reports can be completed.

Within the Information Technology (IT) and GIS work groups, the quality of metadata can affect the duration of system “down times,” the ability to back-up or inter-change IT and GIS analysts, and the time to complete new systems or reporting projects,

### **11.1.3      *Purpose***

This section describes the benefits of metadata, why transit management should care about metadata, an overview of standards and tools, some basic guidelines for how to set up an effective metadata system and some transit agencies examples.

## **11.2    Benefits of Metadata**

### **11.2.1      *Benefits***

The availability of metadata for a transit agency’s datasets is vital for data management, data use, data distribution, and data sharing. Metadata provides a means for end users, internal and external to the transit agency, to determine what transit data are available for a geographic location, to evaluate the fitness of the data for a particulate use or application, to find and access the data, and to transform and process the data.

Metadata lets people find information and understand important things about information packages. Imagine a file room with no order to the labels on the cabinets, the drawers, or the file folders inside the drawers. Imagine office memoranda without the names of originators and intended receivers, without subject or reference headings, and without dates of origination. Metadata provides the labels for spatial data that help users find what they need.

*Metadata protects the transit agency’s internal investment in data.* Metadata is like a card catalog for the agency’s GIS data library. Through metadata, personnel within the organization can see what data are available, what the data are, and how to obtain it. The risks of losing data or recreating data that already exists can be greatly reduced or eliminated.

*Metadata promotes data sharing.* Metadata makes it easy for people in different departments or partner agencies to find out what data other departments or agencies have available.

*Metadata saves time and money when data analyses are needed.* Metadata provides structure and access to information about datasets so that end-users can find answers to common questions more quickly, thus freeing up analyst time for other work.

*Having metadata can save time and money for data development.* With metadata, information vital to a GIS project, contract, program, etc. is always available. Metadata makes data accessible and promotes confidence in data by providing the necessary information for making informed decisions about data use. Metadata helps minimize the delays and costs incurred when data are lost or recreated.

*Metadata promotes data integration.* Metadata can provide information for data exchange. Metadata can serve as interface specifications that include the purpose and “fitness for use” of the dataset.

*Metadata helps ensure appropriate analyses and findings.* Metadata is critical to ensure that the appropriate data are extracted and/or aggregated to produce valid and accurate analyses. For example, knowing the collection date, intended purpose, and accuracy of the spatial data being used for an analysis is absolutely necessary.

### **11.2.2 Why Should Transit Management Care About Metadata?**

Metadata is one of the most critical success factors for accessing and sharing accurate information in a timely manner.

It is the responsibility of transit managers to ensure that the metadata requirements of their organization are defined, and that staff have the time and resources to create and maintain the metadata. An ad hoc approach to metadata, where only some datasets or some operating departments use metadata, is not the most effective approach. Metadata is most useful when the same metadata system is used for all the spatial data in an agency. This does not mean necessarily that all metadata must be done by some centralized department, but it does mean that if different groups are responsible for metadata, they must use the same system.

Metadata may be thought as an institutional policy initiative. Within an organization, the role of metadata management belongs where other information management is being done, such as within an information systems unit that does agency-wide information management work. Some GIS managers and analysts see metadata as inconvenient, and are reluctant to allocate time or resources to create and manage it. If every group does its own information management work, then the agency will need detailed metadata policies, standards and procedures to ensure quality, ease of data access and interoperability.

Metadata costs money but has significant benefits in terms of reduced costs for data sharing, data use, and data maintenance. There are many reasons why a transit agency should maintain metadata on its spatial datasets.

*Metadata is a sound solution for data management.* It provides a consistent way to organize information about data and makes that information about data readily available to others who need it.

*Metadata is a reliable long-term solution.* Without metadata, the agency depends on employees to keep track of data in whatever way they want. Sometimes one person is responsible for storing all the vital information in their head or ad-hoc notes. Since employees come and go, so will knowledge of agency's spatial data. Asking the most long-tenured employee in the GIS department what the original collection method and purpose was for a spatial dataset is not a good management practice. Having a standardized, agency-wide system for recording and maintaining information about datasets is the more reliable and effective solution.

*The most recent version of the National ITS Architecture includes an Archived Data Management Systems (ADMS) user service.* One of the guiding principles of establishing an ADMS is the development of metadata. In a typical regional architecture, the Archive Data Services is a key subsystem. The Federal Transit Administration's (FTA) National ITS Architecture Policy on Transit Projects requires that new ITS projects be consistent with the provisions of the regional architecture. While it is possible to archive data without labeling it for future retrieval, that would not be a best practice, as the main point of archiving the data is to make it available for future use. So if your regional architecture includes the archive data management service, that means your agency will almost certainly need metadata in order to exchange historical information with other stakeholders in the region. Since much of transit archive data is spatially-based, metadata about the base map, transit features and key attributes associated with transit features will be needed in order to deploy this user service effectively.

*It's the law.* An executive order signed by President Clinton on April 11, 1994, mandates that all federal agencies use the Federal Geographic Data Committee (FGDC) metadata standard to document newly-created spatial data. The mandate applies to any organization that receives federal funds to create spatial data, which can include federal, state and local governments as well as private organizations using federal grant money.

Metadata can keep a transit agency out of trouble with the law, yet at the same time, it can create additional evidentiary obligations. Courts and regulatory agencies have treated metadata equally with the information being described, and require that metadata be preserved together with the information being described. There is a growing body of law that gives third parties access to the metadata created and used operationally within systems, and that imposes evidentiary obligations for that metadata upon system managers—obligations that they previously had only for the information itself.

*The alternative to metadata management is information chaos.*

*Example 1.* You are a planner who is new to the transit agency. You are searching in the Automated Passenger Counting (APC) database for bus stops with low ridership that can be eliminated to speed up transit service travel times. You can quickly go to the metadata to find the datasets containing ridership for the bus stops and during the time period of the study. The



metadata also quickly tells you the datasets containing route patterns in effect during the study period. Without the metadata to help, you might be unaware of a temporary route diversion and present highly misleading results on “low ridership” stops. Some of the stops may actually have normally high ridership, but the route diversion caused them to be temporarily skipped, which artificially lowered their normal ridership counts.

*Example 2.* You are viewing some spatial data sent to you from the County GIS Department for your construction project to build a High Occupancy Vehicle (HOV) “jump queue” lane. While reading the metadata that accompanied the spatial data you notice that the data creator has stated in the Purpose section that the data are intended for planning purposes and are at a scale of 1:1,000,000. In this example, without the metadata, the GIS user could have been using low resolution planning data for engineering purposes that require much higher resolutions.

### 11.3 Standards for Transit Metadata

The Federal Geographic Data Committee (FGDC) is the U.S. federal committee responsible for creating and maintaining the most widely adopted standard for metadata currently in use—the FGDC *Content Standards for Digital Geospatial Metadata*, commonly called the FGDC standard. (See <http://www.fgdc.gov/metadata/metadata.html>).

The purpose of the FGDC standard is to provide a common definition for metadata for everyone in the GIS community to follow. By using the common standard, it becomes easier for those within and among different organizations to access data described through metadata. The information included in the FGDC standard was selected based on four roles that metadata play (from FGDC-STD-001-1998):

- Availability—data needed to determine the sets of data that exist for a geographic location.
- Fitness for Use—data needed to determine if a set of data meets a specific need.
- Access—data needed to acquire an identified set of data.
- Transfer—data needed to process and use a set of data.

The American Society for Testing and Materials (ASTM) developed a standard to meet the Archive Data Management User Service (ADUS) requirements of the National ITS Architecture. The standard is intended for key transportation data, such as operations, performance, asset, and government reporting data. Similar to the FGDC standard, the *Standard Guideline for Archiving and Retrieving ITS-Generated Data* identifies three types of metadata information:

- “**Archive structure metadata**—descriptive data about the structure of the data archive itself and of the data and information in the archive that facilitate use of the archive”
- “**Processing documentation metadata** – information that describes the processes applied to data from original source data through to storage in an ADMS”
- “**Data collection system metadata**—data about the conditions and procedures under which original source data were observed, surveyed, measured, gathered, or collected as well as about the equipment that was used.” [ASTM E17.540.01.1 *Standard Guide for Archiving and Retrieving ITS-Generated Data*, 1/9/03, Version 1.9, p. 29]

The data collection metadata requirement in the ASTM ADUS standard is more extensive than the FGDC standard. To meet ITS ADUS requirements, it may be necessary to collect metadata for features and performance measures associated with Automated Vehicle Location (AVL) systems or other mobile data collection units. It includes categories of data such as:

- Initial collection sources—equipment (type, make, model), identification of detectors
- Equipment location and identification—location of fixed equipment and sensors
- Collection conditions—physical environment in which the equipment needed to work. May include weather, special events, time-stamp of actual data collection/observations which may differ from center's time-stamp
- Data collection equipment status and log—working status of equipment used to collect data, and “events” associated with the equipment that may be logged in an operational file
- Equipment self-diagnostics—internal data collection device self test
- Equipment-assigned edit/quality check—error flags assigned by the data collection equipment, and
- Equipment calibration—data such as calibration date, method, equipment, algorithm, field level processing, and results.

Navigation (AVL) platforms and equipment such as global positioning system (GPS), compass, gyroscopes, and odometers are subject to these data collection metadata. Collecting this data can help your agency evaluate the accuracy and suitability for use of a dataset long after the dataset has been collected.

## **11.4 What Information Belongs in a Metadata Record?**

### **11.4.1 Metadata Record Checklist**

Metadata commonly takes the form of a report that is distributed with a dataset, either on CD, on the Internet or an organization's Intranet. The FGDC standard is a style guide that defines what information belongs in a metadata record and the order in which it is presented. Under the FGDC standard, there is a one-to-one relationship between a metadata record and a map layer. In other words, there is one metadata record per map layer or feature class. The FGDC standard defines the following seven sections for each metadata record.

- **Identification Information:** Basic information about where the data originated, how current it is, for what purpose it was created and what geographic area it covers.
- **Data Quality Information:** A data quality assessment describing the accuracy of the data, what steps were followed to create the data, and the source data used to create the data.
- **Spatial Data Organization Information:** Describes the spatial data representation (point, vector, raster) of the feature classes and the relationships that exist between features.
- **Spatial Reference Information:** The description of the reference frame for, and the means to encode, coordinates in the dataset. Examples include the name of and parameters for map projections or grid coordinate systems, horizontal datum, vertical datum, and the coordinate system resolution.

- **Entity and Attribute Information:** Details about the information content of the dataset, including the entity types, their attributes, and the domains from which attribute values may be assigned.
- **Distribution Information:** The process for obtaining the spatial data being described including available media format, size of the dataset, download accessibility and cost.
- **Metadata Reference Information:** Information about the metadata including who is the responsible party, when it was created, when it was last updated, who to contact with questions, and how to contact them.

According to the FGDC standard, only the Identification and Metadata Reference Information sections are required. The other components are “mandatory if applicable.” This is because all spatial data do not have the same characteristics, and all details of datasets being developed or planned may not be decided. Also, the primary purpose of the FGDC standard was to facilitate sharing of datasets between organizations, so only the parts that are absolutely needed for that use are mandatory.

Each section contains individual metadata elements or metadata fields. The FGDC standard divides these elements into three categories: mandatory, mandatory if applicable, and optional. A metadata record that contains all the mandatory sections and elements is considered to be FGDC-compliant.

#### **11.4.2      *Tools for Metadata Management***

Metadata can make information sharing and storage efforts great successes. However, if the metadata itself is not maintained effectively, it can cause errors and embarrassment. For example, if information about the currency and accuracy of a dataset is itself inaccurate or out-of-date, that could be a problem in effectively using or sharing the data. To get the return on investment from metadata, agencies have to get their metadata to work for the people who need it.

Metadata has to be organized and managed. It is managed through people who understand it, and who have and know how to use the tools of metadata management. Among those tools are metadata registers, dictionaries, directories, locators, and search engines. Metadata management is a key part of information management, especially if it is for the purpose of information sharing. Fortunately, many common GIS programs now include a metadata tool.

Examples of some metadata management tools are listed in the Resources section at the end of this section. The metadata tools described in this section are listed for illustrative purposes only and not as specifically recommended products.

*Example:* Intergraph SMMS™ (Spatial Metadata Management System) is an example of a tool for managing metadata that describes spatial data. SMMS is a tool for creating, editing, and querying metadata records. (See <http://www.intergraph.com/gis/smms/default.asp>.) Each SMMS record describes a single map layer (feature class), and is divided into the seven sections described in the FGDC standard. The SMMS reports are in HTML format, which is a common format for metadata dissemination.

### **11.4.3      *Priorities for Implementing Metadata for the First Time***

Metadata development is time-consuming and often overlooked. Analysts often find opportunities to avoid documenting and recording changes to their datasets. If the thought of implementing metadata is overwhelming, this section discusses a minimal set of metadata that should be implemented as an initial stage, and an ordered set of priorities of metadata elements that may be developed over time. This minimal set does not qualify as meeting the requirements of the Regional Architecture and Guidelines set forth by the ADUS standard. However, it does provide a “quick” solution to capturing key information for overwhelmed GIS and IT staff.

The five most important metadata items are:

- Identification of who made the changes, additions or deletions
- Date an update was made
- Description of each field (data element) contained in the database
- Spatial reference information (particularly if more than one location referencing method is used)
- General level of data accuracy and any special considerations (e.g., differing levels of accuracy in a city versus the suburban or rural areas)

The first two metadata elements (above) may be institutionalized into procedures through database configuration management tools.

The next highest priority set of metadata elements are:

- Other identification information including
  - Data purpose and scope
  - Data ownership/stewardship (within the transit agency) information—this is important for internal access but may not be necessary for public distribution
- Accuracy and test procedures for accuracy (including lineage, completeness, consistency, etc.)
- Access and distribution information
- Data collection procedures, including discussion of data collection tools and their quality
- Spatial data organization information

### **11.4.4      *Finding and Identifying Historical Data***

In discussions with transit agencies, desires for historical information are often expressed. Transit agency analysts want and need to know how to access a previous version of a data layer.

When changes are made to a dataset, the older version of the dataset should not be discarded. In particular, the transit feature data and their relation to the base map are needed for historical comparisons, safety and security investigations and many other queries. At a minimum, most transit agencies with ITS applications maintain the prior service change data (transit features and associated base map), the current service change information and a future version that may be in the planning stages. Having a process for recording the dates and special circumstances of these datasets is very important.

In summary, although this Guidebook does not discuss archiving or metadata for archiving, data identification information is important when archiving older versions of base map data, routing, scheduling, ridership, facilities and asset information (with spatial representation). Transit feature data should always be linked to the appropriate base map version(s).

## 11.5 Transit Agency Examples

Two transit agencies are listed below that provide metadata to a wide range of users. These examples illustrate both why they use metadata and how they use spatial metadata.

**Fairfax County, Virginia.** The Central GIS Department for Fairfax County, VA developed an ASP (Active Server Pages) metadata web page application to publish information about the GIS data layers that are publicly available from the central office. See <http://www.fairfaxcounty.gov/maps/map.htm> to view the tabular listing of available data layers. The main motivations for the web application were to:

1. Reduce the amount of employee time spent answering routine questions about the county's datasets (e.g., What are the attributes? When was it last updated? What is the format?), and
2. Reduce the cost to distribute the free data.

Prior to the metadata web page, the GIS staff answered telephone inquiries, determined the customer's data needs, created a CD-ROM, and mailed it to the customer. The new system allows customers to go online to see what is available and download data themselves. This will hopefully allow the agency to reduce costs for data distribution, improve staff morale, and improve service to the public.

The Fairfax County metadata application dynamically loads metadata stored in a database table. This means that the metadata web page automatically reads the database to display a list of all currently available data layers along with the latest associated metadata. The dynamic fields include list of attributes, latest update, notes, and the shape file. The county GIS staff updates the metadata database as regular data files are updated.

The Fairfax County metadata web page also includes an FGDC-compliant metadata report for each data layer. The county uses SMMS to guide the data owners through the process of entering metadata according to the FGDC format and to create the FGDC metadata reports that are included in the county's metadata web page. Unlike the other fields on the county's metadata web page, the SMMS FGDC reports are updated annually. According to Brendan Ford at Fairfax County, "Implementing a metadata management tool refines and makes the metadata management process easier. SMMS makes the process easy for the user to enter and update the metadata. However, the software is rigid in guiding users through the process, and the online help could be more effective." Metadata reporting with SMMS is limited to the FGDC standard record for each data layer, which may not fit the information requirements of some users.

Fairfax County makes the spatial dataset metadata reports accessible by putting them on the county's web pages in the list of data available from the County. There have been some issues with getting data providers to create metadata reports. To help address this, the County designated a metadata coordinator who helps data providers track down information and answer questions regarding the metadata. It is too early to qualify the success of the county's metadata management effort or to quantify the pay off. As a first step, the agency is trying to get metrics on how often people access metadata on the web.

**Orange County Transportation Authority (OCTA).** OCTA uses ESRI's ArcCatalog metadata environment to document and maintain metadata on their GIS data and projects. According to Shirley Hsiao, "At OCTA, we maintain over 300 datasets and projects on a regular basis. It is vital to our organization that we have a document management system that is accessible, useable, and stable, and meets the Federal Geographic Data Committee's (FDGC) metadata requirements." On a daily basis, OCTA uses ArcCatalog to document the master GIS data as well as completed GIS projects. Ms. Hsiao says, "ArcCatalog provides staff with tools to easily view, create, update, search, or print our GIS data and metadata. Users can search by keywords, view and print attribute definitions, as well as update metadata with current dates, additions, or changes. In regards to distribution, the metadata can be exported into HTML or XML formats that can be easily viewed by staff and other agencies using an Internet browser."

The OCTA manages both general content metadata needed to select and evaluate datasets, and specific metadata needed to integrate datasets. Their total "Master GIS" metadata is partitioned it into two parts:

1. Clearinghouse information. Content metadata is of interest to the general users. It includes items such as who created the data, update date, data dictionary, etc.
2. Project directory. Integration metadata indicates how projects currently integrate data and supports new GIS project analyses. It provides information both on how data are currently integrated in projects, and how data could be integrated for different analyses. It gives sources of data, information on how data are integrated, and procedures for combining data. This reference documentation is useful when multiple team members update or modify complicated spatial analyses.

## **11.6 Think Beyond**

### **11.6.1      *Quality and Procedures for Measuring Accuracy***

Metadata records can contain fields for expressing the accuracy of data. However, without standard procedures for measuring accuracy, this metadata field has limited value. Agencies want quality tests/methods for attribute accuracy as well as spatial accuracy. Transit agency analysts want and need to know the origins of derived data as well. Section 7 describes concepts and practices for assessing and documenting spatial data quality.

### 11.6.2 *Integration Metadata*

Metadata can serve an agency by indexing and cataloging the datasets for integration, aggregation and processing. Metadata standards, such as FGDC's Content Standard for Digital Geospatial Metadata, focus on the "content metadata." Another type of metadata, "integration metadata," describes data sent to and from independently developed application systems (Schulte et al. 2002). Integration metadata contains information about the communication mechanisms and content between applications. The integration metadata includes the touchpoint identities of the sender and receiver applications, and the interaction process mechanics and business implications. An example of a touchpoint is an Interface Requirement Document (IRD) contract between application software that establishes a registry of common data items. The useful kinds of application integration metadata, specialized tools and languages, and strategies for managing integration metadata are described in Section 9.

## 11.7 Resources and References

### References on Metadata

"What Every CIO Needs to Know About Metadata". *Federal Government CIO Council's Interoperability Committee*. February, 1999

[http://www.cio.gov/documents/metadata\\_cio\\_knowledge\\_feb\\_1999.html](http://www.cio.gov/documents/metadata_cio_knowledge_feb_1999.html)

Schulte, R., M. Blecher, J. Thompson, W. Rishel, and M. Pezzini. (2002a). "Strategies for Managing Application Integration Metadata," Research Note T-15-5501. Gartner, Inc. March.

Schulte, R., W. Rishel, J. Thompson, and J. Sinur. (2002b). "Identifying Useful Kinds of Application Integration Metadata," Research Note TU-15-1931. Gartner, Inc. March.

Schulte, R. J. Thompson, M. Pezzini, W. Rishel, and M. Blechar. (2002c). "Tools and Sources for Application Integration Metadata," Research Note T-15-1932. Gartner, Inc. March.

Schulte, R., J. Thompson, W. Rishel, and B. Lheureux. (2002d). "Best Practices: Managing Application Integration Metadata,". Research Note TG-15-1933. Gartner, Inc. March.

### Metadata Standards

Federal Geographic Data Committee. (1998). *Content Standard for Digital Geospatial Metadata*. FGDC-STD-001-1998.

[http://www.fgdc.gov/standards/documents/standards/metadata/v2\\_0698.pdf](http://www.fgdc.gov/standards/documents/standards/metadata/v2_0698.pdf)

ASTM E2259-03 Standard Guide for Archiving and Retrieving ITS-Generated Data. Advisory includes an overview of several projects that demonstrate how agencies are using the standard to develop and implement Archived Data Management Systems (ADMS). The Advisory can be found at:

[http://www.its-standards.net/Documents/ADUS\\_Advisory.pdf](http://www.its-standards.net/Documents/ADUS_Advisory.pdf) (PDF)

[http://www.its-standards.net/Documents/ADUS\\_advisory.htm](http://www.its-standards.net/Documents/ADUS_advisory.htm) (HTML)

### **Metadata Tools: How to Choose?**

Users selecting a metadata management tool for their working environment are faced with a confusing array of choices. As a starting point, Table 11-1 lists metadata tools for some of the most common GIS working environments. The metadata tools produce metadata according to the 1998 Content Standards for Digital Geospatial Metadata (CSDGM).

This webpage <http://badger.state.wi.us/agencies/wlib/sco/metatool/> leads to summaries of the metadata management tools shown in Table 11-1 for documenting spatial data and supporting spatial metadata. Most of the tools were designed to manage CSDGM metadata.

**Table 11-1. Metadata Management Tools for Common GIS Environments**

[Hugh Phillips, June 8, 2002, <http://badger.state.wi.us/agencies/wlib/sco/metatool/toolchoz.htm>]

Environment	Metadata editor	Metadata Utilities
ArcGIS 8.0+, ArcView 8.1+	<a href="#">ArcCatalog</a> (FGDC Metadata Editor)	<a href="#">cns</a> and <a href="#">mp</a> may be useful to prepare outside metadata for import into ArcCatalog. <a href="#">mp</a> may be preferred to produce HTML output over ArcCatalog because it can be customized. Confident metadata editors may find that using <a href="#">tkme</a> to edit the ArcCatalog XML files affords advantage for certain editing operations.
ArcInfo < 8.0	<a href="#">tkme</a> (or <a href="#">xtme</a> with UNIX platform ArcInfo)	Use <a href="#">FGDCMETA (AML)</a> to extract ArcInfo dataset 'property' information, then import that into <a href="#">tkme</a> . Use <a href="#">mp</a> to validate metadata and convert it to output forms including XML, SGML and HTML.
ArcInfo < 8.0, Arcview 3.x	<a href="#">SMMS</a>	<a href="#">cns</a> and <a href="#">mp</a> may be useful to prepare outside metadata for import into SMMS. <a href="#">mp</a> may be preferred to produce HTML output over SMMS because it can be customized.
ArcView 3.x	<a href="#">tome</a>	<a href="#">DBFmeta</a> should be used to create Entity and Attribute metadata snippet. Use <a href="#">mp</a> to validate metadata and convert it to output forms including XML, SGML and HTML.
Geomedia	<a href="#">SMMS for Geomedia</a>	
Other GIS environments	<a href="#">tkme</a> or <a href="#">SMMS</a>	both of these tools can operate independent of any particular GIS software. Use <a href="#">mp</a> to validate metadata and format it for output



## **APPENDIX A    GUIDEBOOK DEVELOPMENT PROCESS**

The Guidebook was developed in Phase 2 of the Location Referencing Project over a period of 24 months through the Transit Standards Consortium (TSC) under U.S. Government Cooperative Agreement FTA C-6, October 1, 1999.

The process for identifying the best practices involved the following five main data collection efforts:

- A Feasibility Study in Phase I of the LRG Project resulting in Technical Memorandum #1
- A Review of Issues Related to Existing and Emerging Spatial Data Interoperability Standards (Technical Memorandum #2) developed in Phase I of the LRG Project
- Interviews of a diverse set of transit professionals
- Transit industry teleconferences
- Literature review and solicitation of transit solutions and documents

### **Interviews**

First, the authors conducted 11 interviews with analysts and GIS managers at transit agencies and regional authorities known to be leaders in the area of spatial data management. These included small and large, regional and local transit agencies. The purpose of these interviews was to gather spatial data sharing and location referencing issues in the transit industry and to identify at a high level, some existing data policies and best practices. The following are examples of some of the questions that were asked in the interviews.

- What ITS systems or major applications in your agency are supported by map databases (e.g. AVL, Demand Responsive Reservation and Dispatch, Scheduling, APC, Itinerary Planning, etc.)?
- Describe the extent of spatial data sharing between your ITS systems and applications.
- How many map databases does your agency support?
- What location referencing methods does your agency support (GPS coordinate, bus stop-offset, etc.)?
- How does your agency maintain the street network or obtain an updated version of it (data source, frequency of update, update procedure, etc.)?
- How does the agency maintain the route network and bus stop locations (data source, frequency of update, update procedure, etc.)?
- How is the route network topology connected to the cartography (maps)?
- How does the agency ensure consistency among the different map datasets?
- What are your biggest challenges in the spatial data update process for all the applications that use spatial data?
- What are or have been the biggest barriers to exchanging spatial data between systems within your organization?
- What are or have been the biggest barriers to exchanging spatial data with other organizations?
- Focusing just on location referencing issues, what have been the biggest barriers to spatial data exchange?

- Assuming your wishes could be fulfilled, what tools or information standards would be the most helpful to solve your problems with spatial data exchange?
- What are the best “lessons learned”, tools or tricks that helped you solve your problems (that might be shared with another transit agency)?
- When do you have to deal with temporal aspects of spatial data?
- How do you manage time attributes in the GIS?
- What are your biggest challenges to sharing dynamic spatial data and data with temporal attributes? (e.g. real-time updates of AVL for schedule adherence, event triggers, time of arrivals, or signal priority data and e.g., trips, service patterns, and performance data, respectively)?
- Does your agency have an enterprise architecture or use enterprise information strategies for spatial data?
- Does your agency maintain a policy or standard for data interchange, metadata or data quality?
- What are your strategies (during procurement, development, operations and maintenance phases) for ensuring successful exchanges of spatial data between systems?

## **Workshops**

The authors facilitated ten teleconference workshops that were hosted by the TSC. The results of the interviews described above were synthesized and used to develop the workshop topics shown in Table A-1). The workshops were advertised, plus speakers and participants were invited. The workshop participants included subject matter experts from transit agencies, counties, GIS software vendors, commercial map database vendors, etc.

Most of the workshops included short presentations from one or two of the invited participants followed by facilitated group discussion with an average of 11 participants. Forty-six individuals participated in the workshop teleconferences. Over half of the individuals participated in more than one workshop. The teleconference workshops were scheduled for 90 minutes.

This Guidebook is the product of a formal review process that included posting the draft Guidebook for industry review and comment on the project website, four teleconferences for receiving comments, voice and email feedback and review by the Federal Transit Administration.

**Table A-1. Guidebook Workshops**

#	Workshop Name	Track	Date (2002)
1	Acquiring and Maintaining the Base Map: Trade-offs and Life Cycle Issues	Base Map	Oct 23
2	“Single Base Map”: what, why and how?	Base Map / Management	Oct 29
3	Strategies for Managing AVL/APC Data	Transit Data	Nov 1
4	Data Quality and Metadata: What are the Essentials?	Transit Data	Nov 5
5	Managing Relationships of Transit Feature data	Transit Data	Nov 6
6	Intra- and Inter-agency Coordination and	Management	Nov 12

#	Workshop Name	Track	Date (2002)
	Dependencies: How to make it work		
7	Strategies for Successfully Procuring IT/ITS/GIS Solutions: Potential Pitfalls when Implementing a Proprietary System	Management	Nov 15
8	“Roadmap” for managers on achieving an Enterprise Spatial Data Infrastructure	Management	Nov 19
9	Guidelines by ITS Applications	General	Nov 20
10	Vision and Next Steps for the Transit Business: Message to the Vendor	General	Dec 6

### **Other Sources of Information**

The last source of information for deriving best practices includes personal communications as follow-up to the workshops, published literature, and internal documents from transit and county agencies.

## **APPENDIX B    RECOMMENDATIONS FOR BASE MAP ATTRIBUTES**

In this appendix, we discuss some of the critical attributes that should be included in the base map as well as identify attributes that support key transit applications. We also list attributes that support analysis and advanced applications that may support key business goals.

### **Critical Attributes**

- Centerline and topology
- Street names, address ranges, Unique ID, ZIP code, direction (one- or two-way)
- Street name alias table
- Road restrictions (height, weight, direction of travel)
- Road classification, speed limits and number of lanes
- Political boundaries
- Transfer points to other modes (e.g. rail service transfer points)

### **Extra Attributes**

- Other modal networks
  - Walking paths with obstacles to walking and curb cut locations
  - Bicycle paths
  - Rail alignment and tracks
- Turn impedances (e.g., turn tables)
- For ADA compliance: sidewalk locations, dimensions and condition; locations of news stands, vending machines, fire hydrants, parking meters, mailboxes, trash receptacles, utility poles, trees, and traffic signs
- Grade
- Catchment zones
- Landmarks
- Demographics and land use by political boundary

## APPENDIX C TCIP LOCATION REFERENCING METHODS

This table defines how spatial elements are defined in TCIP. If you are exchanging data between applications using TCIP-conformant messages, you will need to use these definitions to ensure format interchangeability.

**Table C-1. TCIP Location Referencing Methods**

Location Class	Location Referencing Method	Definition [from NTCIP 1405:2000]
Point	SpAddresspoint	A point defined by an address. The addressing format is based on the USPS conventions.
	SpNodeOffsetpoint	A point defined by an offset from a node.
	SpNodePercentOffsetpoint	A point defined by the normalized (0.00 to 100.00) distance between two nodes. The offset is constrained to be a number between 0 to 100, and it is measured from the first node to the second node.
	SpSPpoint	A point defined by state plane coordinates and zone identifier. The datum is assumed to be NAD83. Coordinates based on NAD27 must include the SpDatum data element.
	SpRoadLabelpoint	A name or attribute associated with a point on a map. This is a cartographic feature of a map.
	SpGeoDynamicpoint	A dynamic point expressed by the magnitude of the path of a moving object.
	SpGeopoint	A point defined by geographic coordinates. The datum is assumed to be NAD83 if no other datum is specified.
	SpGeoLpoint	A point defined by geographic coordinates and label. This entity may be used as a label placement for any type of geographic feature.
	SpGeoOffsetpoint	A point defined by an offset from a geographic point.
	SpIntpoint	A point defined by the junction of two or more roads.
	SpIntOffsetpoint	A point defined by an offset from an intersection.
	SpLandmarkpoint	A location defined by a known point of interest such as a place, building, square or object.
	SpMilepostpoint	A point defined by a milepost marker or an offset from a milepost marker.
	SpNodepoint	A point defined by a node identifier.
Line	SpAddressRangeline	A line defined by two road numbers and the road segment between them.
	SpGeoline	A line defined by a set of geographic points. If more than two points are specified, the first and last points in the sequence are assumed to be the termini and the other points are shape points.
	SpGeoLline	A line defined by a set of geographic labeled points.
	SpIntline	A line defined by a set of intersections
	SpIntOffsetline	A line defined by points which are referenced to (offset from) intersections, e.g., the line between two bus stops.
	SpLinkline	A line defined by a unique link identifier.
	SpMilepostline	A line defined by a set of mileposts.
	SpNodeLinkline	A link defined by a pair of node identifiers.
	SpSpline	A line defined by a set of state plane coordinates. If more than two coordinates are included, then the first and last pairs are assumed to be the start and end points (respectively), and the other points describe the shape points of the "polyline" or cartographic segment.
Polygon	SpCentroidpolygon	The centroid defines the center of a polygon

<b>Location Class</b>	<b>Location Referencing Method</b>	<b>Definition [from NTCIP 1405:2000]</b>
	SpP_Geopolygon	A polygon defined by geographic points.
	SpP_Intpolygon	A polygon defined by an ordered sequence of intersections. This structure may be used to define a fare zone or service area, for example.
	SpP_Nodepolygon	A polygon defined by node identifiers. Examples of use include planning applications.
	SpP_Sppolygon	A polygon specified by geographic coordinates as defined by the State Plane coordinate system.
	SpL_AddressRangepolygon	A polygon defined by a sequence of address ranges. For example a zip code polygon as defined by the address ranges enclosing the area.
	SpL_GeoPolygon	A polygon defined by geographic labeled lines.
	SpL_Linkpolygon	A polygon defined by a sequence of consecutive, connected links.
Route	SpP_Addressroute	A route defined by a sequence of addresses. For example, a manifest or list of pickup/dropoff points for demand responsive service.
	SpP_Georoute	A route defined by geographic points.
	SpP_Introute	A route defined by a sequence of intersections. This structure may be used to define a trip (SchTrip) identified by time points or driving directions based on turning locations, for example.
	SpP_IntOffsetroute	A route defined by an offset from an intersection. Only those points that are sufficient to distinguish the shape of the route along a linear network need be included. This structure may be used to describe a pattern (SchPattern) from bus stop (CptStopPoint) to bus stop, for example.
	SpP_Milepostroute	A route defined by mileposts or distance markers along a route. For example, events (assets or incidents) on state and national highways are typically recorded by milepost or mile markers. Transit routes may be defined by distance from origin; this structure can be used to define the segments of a transit route.
	SpP_Noderoute	A route defined by a sequence of node identifiers.
	SpP_SProute	A route specified by geographic points as defined by the State Plane Coordinate System.
	SpL_GeoLroute	A route defined by a sequence of geographic labeled lines.
	SpL_Linkroute	A route defined by a sequence of links.
Other	SpLRMSWrapper	A message packet to insert a Location Referencing Message Specification [SAE J2374, Rev C (12/97)].

## **APPENDIX D THE GEOSPATIAL ONE-STOP'S TRANSIT REFERENCE MODEL**

Portions of this Appendix are adapted from the “Geographic Information Framework—Data Content Standards for Transportation Networks: Transit”

The primary purpose of the Data Content Standard for transit standards is to support the exchange of transportation data related to transit systems. The standard also seeks to establish a common baseline for the content of transit databases for public agencies and private enterprises. It seeks to decrease the costs of acquiring and exchanging transit data for local, tribal, state, federal, and North American users and creators of transit data. Benefits of adopting the standard also include the long-term improvement of the geospatial transportation base data, improved integration of safety, emergency response, and enforcement data, and streamlined maintenance procedures. Since transit usually shares data with other transportation, planning and emergency services organizations and acquires data from county or state planning organizations, seamless integration with other datasets is critical.

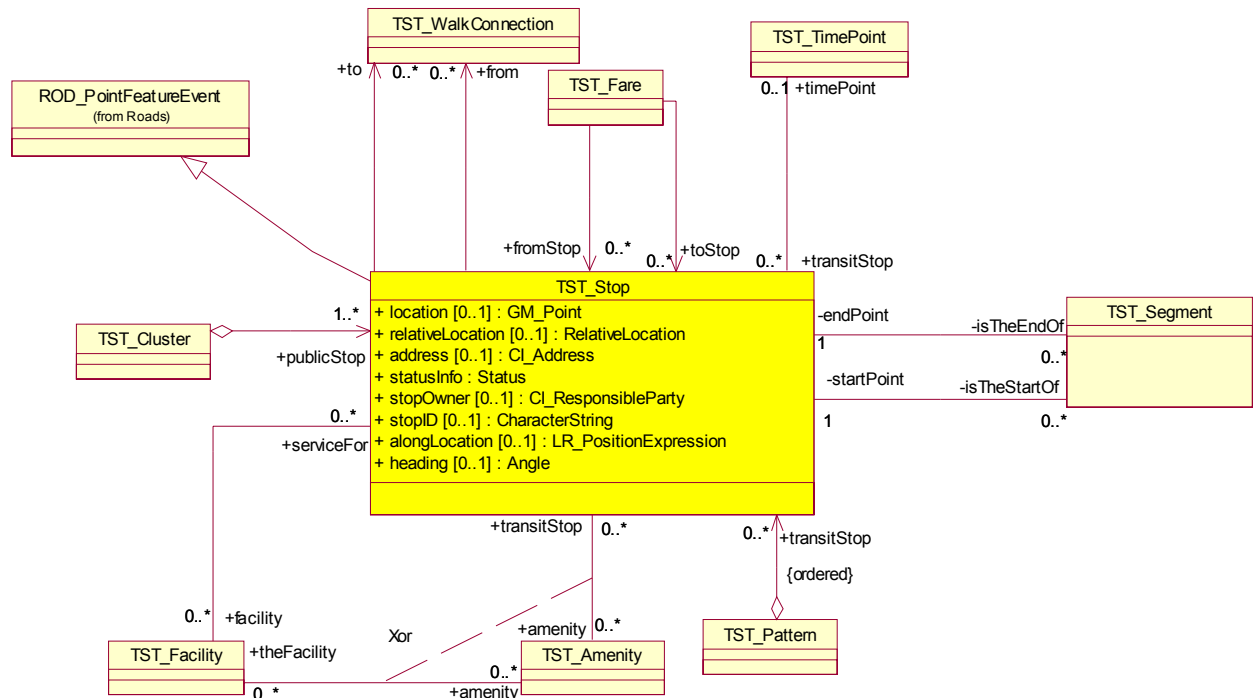
The Content Standard was developed to fulfill one of the objectives of the National Spatial Data Infrastructure (NSDI) which is to create common geographic base data for seven critical data themes. The core themes are considered Framework data, reflecting their critical importance as geographic infrastructure. The Geospatial One-Stop (GOS) initiative is an e-government initiative of the federal government designed to expedite the creation of the seven Framework layers.

The “Geographic Information Framework—Data Content Standards for Transportation Networks: Transit” standard (hereafter referred to as GOS-Transit) defines components of public transportation systems. Transit is one of at least five modes that compose the transportation theme of the GOS data framework. As a transportation mode, transit differs substantially from other modes such as rail and road in that the core operational features of the transit system consist of spatial and temporal elements that rest upon the transportation infrastructure. The transit system described in GOS-Transit is made up of linear features (such as roads), the route or pattern features that rest upon linear features, system stop locations, and other features associated with the operation of transit systems. Specifically, GOS-Transit describes a model with three main components:

- “A Segmentation Model inherited from the Road model, that defines segments, collections of segments and their associated geometries and topology.
- A Transit Stop model, which defines a node of activity where transit passengers transfer to or from public transit vehicles, as well as the scheduling and operational information associated with the transit stops. [This model inherits geometry and topology from the Road model and Segmentation model and linear attributes from the Linear Referencing Model.]
- A Linear Reference Model (LRM), which defines a measurement method used apply attributes to segments by locating their endpoints and define their extent.”  
[GOS-Transit, pf., 13]

For example, in Figure D-1, the TST Stop model inherits (note the open arrow notation) from the ROD\_PointFeatureEvent, an attribute that is “attached” to the underlying road network (or base map). TST\_Stop supports the definition of the TST\_Segment, and may be referenced by a position alongLocation of the TST\_Pattern, TST\_Segment, or TST\_WalkConnection.

**Figure D-1: Context Diagram for TST\_Stop [from GOS-Transit, Figure 4, p., 16]**



The location reference semantics of each class are defined with respect to the Linear Reference System and the Road model, thus ensuring a strong association with the transportation network which underlies the transit feature. For example, the semantic description for TST\_Stop states:

“TST\_Stop is the central feature of the Transit model because it conveys positional information as well as temporal information deemed essential to the exchange of transit data. The TST\_Stop is shown in Figure D-2 as a special type of ROD\_PointFeatureEvent. (See the GOS Road Standard for a complete description of the event model). TST\_Stop has the geometry of type GM\_Point as defined in ISO 19107 [7] and its relativeLocation is given by a position expression. Figure D-1 shows that the TST\_Stop is the starting point for one TST\_Segment as well as the end point for a different TST\_Segment. The TST\_Stop may also participate as part of a Pattern, an ordered set of TST\_Stops. [p., 15]

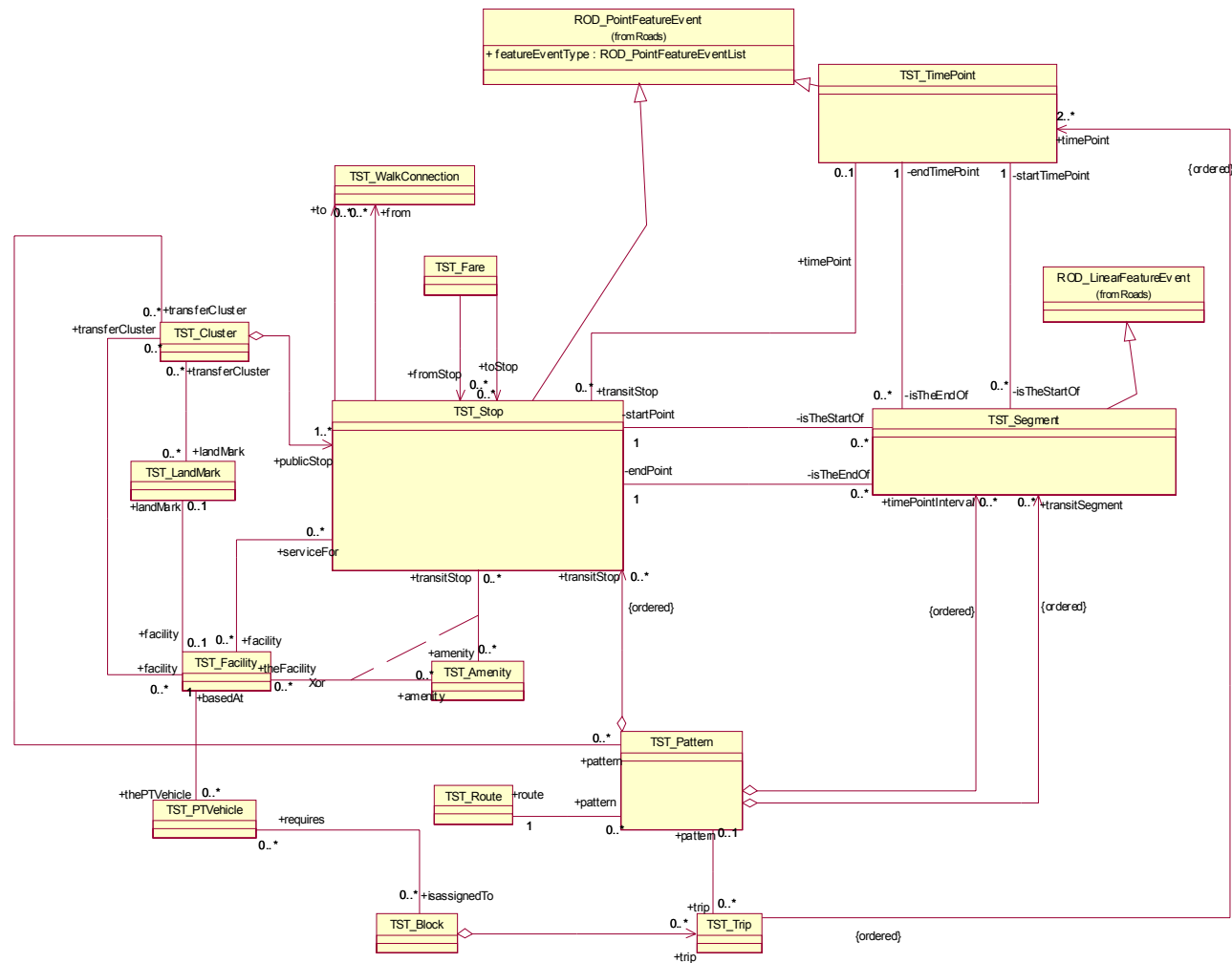
Transit point features are referenced to GM\_Point, transit segments are referenced to GM\_Curve and LR\_PositionExpression (for start and end points); transit polygon features are referenced to GM\_Polygon.



The classes, features and characteristics included in this standard were developed as part of a comprehensive review of several use cases. The use cases support data exchange on a regional level to support itinerary planning, stop sharing, and re-routing applications.

Although these use cases address the operational requirements of a broad selection of Transit business scenarios, they did not provide sufficient input to design an all-inclusive Transit model. Additional use cases will be required to identify the universe of classes, features, and characteristics necessary to fully describe the Transit geographic base data theme. Nevertheless, the core service planning features were defined in sufficient detail to support a majority of core Intelligent Transportation System (ITS) application.

The standard can be implemented using a variety of software packages and is designed to accommodate data encoded without geometry as well as to support the exchange of data encoded in a variety of geographic information systems. It is designed to be able to depict the complete transit system at all levels of service and all functional classes that may be defined by a data-providing agency.



## APPENDIX E ADA COMPLIANCE AND BUS STOP ACCESSIBILITY DECISION TREE

This example from Miami-Dade Transit (MDT) shows the importance of recording the procedures for collecting and applying location information.

[MDT GIS staff provided the tables and decision trees in this Appendix. The materials were created by MDT End-Users and GIS staff.]

**Table E-1: MDT Delivered Data File Formats**

<b>SHP/DBF – Collection Data</b>				
<b>Field Name</b>	<b>Type</b>	<b>Size</b>	<b>Values</b>	<b>Comments</b>
BusStop_ID	Character	13	bsd-nnnn.nnnn	Concatenate District + "-" + Stop_Num (Stop_Num to be padded with leading zeros)
District	Character	3		MDT Bus Stop District
Stop_Num	Integer	9	nnnn.nnnn	Stored without leading zeros
Main_St	Character	28		"Sn" field in street layer shapefile. Size 28 - Always a street/"Street the tires of the bus are on."
Cross_St	Character	28		"Sn" field in street layer shapefile. Size 28 -May be an address/house number
Direction	Character	2	North/South/East/ West/NorthEast/North West/SouthEast/ SouthWest	General direction of bus route
Stop_Cornr	Character	1	Nearside/Farside/ MidBlock/Terminal	Stop location in reference to Cross_St. <ul style="list-style-type: none"> <li>MidBlock stop defined as one where no cross street is near by.</li> <li>Entire terminal area considered one stop for collection purposes.</li> </ul>
ADA_Comp	Character	1	Yes/No	Determined programmatically
Accessible	Character	1	Yes/No	Determined programmatically
X_coord	Fixed	12		From GPS unit w/in +/- 10 ft
Y_coord	Fixed	12		From GPS unit w/in +/- 10 ft
Benches	Integer	2	0-99	Number of benches
Shelters	Integer	1	0-9	Number of shelters
Trash_Can	Integer	1	0-9	Number of trash cans
Sign	Character	1	Yes/No	
Picture	Character	1	Yes/No	Picture exists?
Updte_By	Character	20		Inspector name
Updte_Date	Character	10		mm/dd/yyyy
Updte_Time	Character	5		hh:mm
Comments	Character	100		

NOTE: Shaded fields are to be used as reference only and will not be updated.

Supplemental ADA Compliance Information – ADA\_STOP.DBF

<b>DBF – ADA Compliance Data</b>				
<b>Field Name</b>	<b>Type</b>	<b>Size</b>	<b>Values</b>	<b>Comments</b>
BusStop_ID	Character	13		Key to relate back to shape/dbf
Sidewalk	Character	1	Yes/No	Is there a 36" wide path that extends from Curb_Cut2 to Curb_Cut3?
FSS	Character	1	Yes/No	Is there a 5' x 5' (min), clear, firm, stable surface at the bus stop?
FSS_4ft	Character	2	Yes/No/NA	IF yes FSS, is FSS within 4' of EOP?
FSS_DWAccs	Character	2	Yes/No/NA	Is there a 36" wide access point from driveway to sidewalk?
FSS_EPAccs	Character	2	Yes/No/NA	Is there a 36" wide access point from FSS to EOP?
FSS_SWAccs	Character	2	Yes/No/NA	Is there a 36" wide access point from FSS to sidewalk?
FSS_BusCon	Character	2	Yes/No/NA	IF no FSS_4ft, is there a 36" wide connector from EOP to FSS?
Driveway	Character	2	Yes/No/NA	IF no FSS, is there a hard-surface driveway within 40' of bus stop?
Shltr_Con	Character	2	Yes/No/NA	If Shelter > 0, is there a 36" wide accessible path from shelter to FSS?
Obs_TreeBr	Character	1	Yes/No	Tree branch hangs below 80"
Obs_TreeRt	Character	1	Yes/No	Tree root protrudes greater than 1/2"
Obs_News	Character	1	Yes/No	Newsstand vending machine
Obs_UtPole	Character	1	Yes/No	Utility pole (electrical, telephone, traffic/street light)
Obs_PrkMtr	Character	1	Yes/No	Parking meter
Obs_UtBox	Character	1	Yes/No	Utility box
Obs_Bench	Character	1	Yes/No	Bus bench
Obs_FireHy	Character	1	Yes/No	Fire hydrant
Obs_MailBx	Character	1	Yes/No	Mail box
Obs_Trash	Character	1	Yes/No	Trash receptacle
Obs_TrfSgn	Character	1	Yes/No	Traffic sign
Obs_XSlope	Character	1	Yes/No	Need 5' clearance before severe slope
Obs_SWBrok	Character	1	Yes/No	Broken sidewalk w/protrusion greater than 1/2"
Obs_Other	Character	1	Yes/No	Any other obstruction?
Comment	Character	50		Significant remarks affecting ADA or detail of Obs_Other
Curb_Cut1	Character	2	<b>CC</b> = CurbCut <b>CN</b> = Connector <b>NA</b> = Not Applicable <b>NO</b> = Missing	CC – Area where sidewalk slopes down to EOP CN – Where no curb, connects sidewalk to EOP NA – Situation does not require CC or CN NO - Missing CC or CN where required
Curb_Cut2	Character	2		
Curb_Cut3	Character	2		
Curb_Cut4	Character	2		





## **APPENDIX F    SAMPLE DATA DICTIONARY FOR BUS STOP SIGN**

A data dictionary is a listing of the core transit feature fields supported within your organization. It is assumed that each transit feature is composed of numerous attributes or fields. In the data dictionary example on the following pages, the fields for one transit feature, BUS STOP SIGN, are listed. At a minimum, you should include the following entries that describe each BUS STOP SIGN attribute:

- Name
- Definition
- Data Type
- Owner
- [optional] Units (if applicable)
- [optional] Keyword (from a standard list)
- [optional] Remarks
- [optional] Data Collection and Validation Process (if applicable)
- [optional] Representative Layout or Rule (masking or code list)
- [optional] Last Update
- [optional] Last Editor

Furthermore, if there is an interchange format for the field or feature, then that too should be described. For example, many organizations may describe the interchange format for the BUS STOP SIGN by using the Transit Communications Interface Profile in its XML form. The Simple or Complex Type definitions may be included in the Data Dictionary as another field.

Another dictionary that you will need is the description of your Transit Features. The transit features may be described in the same table as its attributes although another column should be included to differentiate between a Feature and Attribute record.

Name	Definition	Data Type	Owner	Units	Data Collection	Layout/Rules	Remarks	Update
<b>ID</b>	Unique identifier for each stop sign	Long Integer	Fac. Plan					Feb 2003
<b>MarkerType</b>	The type of mounting device sign is mounted on	Enumerated	Fac. Plan			Metal pole; wooden post; shelter	Added “shelter” in April 2003	April 2003
<b>MarkerOwner</b>	The owner of the marker	String	Fac. Plan					Feb 2003
<b>HeightFromBase</b>	The height of the sign from the ground or base of marker.	Real	Fac. Plan	ft	FC-12.doc (includes how the sign is measured from the ground or base of marker)			Feb 2003
<b>AtStreet</b>	The street on which the sign is posted.	String	Fac. Plan		Validate spelling with Network street lookup			Feb 2003
<b>CrossStreet</b>	The nearest cross street to the posted sign	String	Fac. Plan		Validate spelling with Network street lookup			Feb 2003
<b>IntersectionOffset</b>	The distance from the nearest intersection (as defined by the At and Cross Streets) nearside curb to the marker.	Real	Fac. Plan	ft	FC-12.doc			Feb 2003
<b>SetbackDistance</b>	The perpendicular distance from the curb to the marker	Real	Fac. Plan	ft	FC-12.doc			Feb 2003
<b>Latitude</b>	The latitude of the marker position	Long Real	Fac. Plan	Decimal degrees	NAD 83			Feb 2003
<b>Longitude</b>	The longitude of the marker position	Long Real	Fac. Plan	Decimal degrees	NAD 83			Feb 2003
<b>StopPointID</b>	A unique identifier for the stop point which the sign marks.	Long Integer	Fac. Plan		Validate ID with StopPointInventory			Feb 2003
<b>Description</b>	The words written on the sign.	Memo	Fac. Plan					Feb 2003
<b>Graphic</b>	A photograph or graphic of the sign	Blob	Fac. Plan			Store as SignID#.jpg where # = ID		Feb 2003
<b>Dimensions</b>	The size (width and height) of the sign.	Memo	Fac. Plan	inch		Store as [wid by ht]		Feb 2003



Name	Definition	Data Type	Owner	Units	Data Collection	Layout/Rules	Remarks	Update
Quantity	The number of signs for this stop point currently in stock	Integer	Fac. Plan					Feb 2003
Survey Procedures	A description of the survey equipment and procedures used to collect location information. Include field survey crew ID.	Memo	Field Survey Crew		Cross reference field collection sensor serial number with list of sensors.		Refer to FC-10.doc for instructions. List changes, deviations & exceptions.	August 2003
SurveyDate	Date sign was surveyed	Date	Fac. Plan			MM/DD/YYYY		Feb 2003
ActivationDate	Date stop point was activated	Date	Fac. Plan		Validate against StopPointID.ActivationDate	MM/DD/YYYY		Feb 2003
DeactivationDate	Date stop point was deactivated	Date	Fac. Plan		Validate against StopPointID.DeactivationDate	MM/DD/YYYY		Feb 2003
UpdateDBdate	Date inserted into Sign Stop database	Date	IS			MM/DD/YYYY		Feb 2003

## APPENDIX G TRANSIT APPLICATIONS/FUNCTIONS THAT USE SPATIAL DATA AND SPATIAL ANALYSIS FUNCTIONS (GIS TOOLS)

Presented from a high level perspective, the tables included in this section provide transit management, project managers and technical staff with the ability to:

- See a broad range of opportunities for using spatial data to improve transit
- Identify other stakeholders, so agency-wide solutions may be developed
- Leverage technology investments across transit by encouraging collaboration and data sharing

The tables describe transit functions, the spatial analysis/GIS functions performed as part of those transit functions; and for selected ITS applications, the key transit feature data that are necessary to implement those applications.

**Listed by Transit ITS Applications.** Tables 8A-1a through 8A-1d provide examples of automated ITS based applications. The applications are grouped in four broad ITS categories, Transit Management System, Automated Traveler Information System, Electronic Fare Payment and Transportation Demand Management. Table 8A-1a also includes transit features used by the application or function.

**Listed by Transit Business Functions.** Tables 8A-2 through 8A-13 present spatial based applications by transit business area without highlighting ITS applications. By referring back to the Section 8 Table of Contents, the list of transit business functions that are included can be seen.

### **By Transit ITS Application Area**

**Table G1. Transit Management System**

Supporting Task	GIS Application/Function	Transit Feature(s)
Schedule Adherence	<p>Track and display vehicles; color-code vehicle with its adherence status.</p> <p>Match vehicle location/time at expected block/trip time point location and time.</p> <p>Record run times between bus stops.</p> <p>Project run times between time points</p>	<ul style="list-style-type: none"> <li>• Street centerline;</li> <li>• Bus route, vehicle/operator assignment, Ordered sequence of bus stops / trip, ordered sequence of time point/trip;</li> <li>• Trip time point time</li> <li>• Current time/location of transit vehicle per trip;</li> <li>• Bus stop area; bus stop location (Time upon entering and leaving bus stop area)</li> <li>• Historical data on running times between transit segments (e.g., time point intervals, bus stop intervals, patterns, etc.)</li> </ul>

<b>Supporting Task</b>	<b>GIS Application/Function</b>	<b>Transit Feature(s)</b>
Route Adherence	Track and display current vehicle position relative to scheduled pattern, route, block and trip.	<ul style="list-style-type: none"> <li>• Route pattern; vehicle assignment, trip, piece of work (series of ordered time point or bus stop locations)</li> <li>• Current location of transit vehicle;</li> <li>• Street centerline</li> </ul>
Fleet Management	<p>Track and display vehicle on base map.</p> <p>Hold vehicle at transfer stop based on timed transfer data that meets a certain threshold (estimated time of arrival of revenue vehicle with transfer passengers).</p>	<ul style="list-style-type: none"> <li>• Street centerline</li> <li>• Current location/time of transit vehicle per fleet (must include vehicle type)</li> <li>• Expected (scheduled) location of transit vehicles at time (e.g., trip time point times for each trip/route/vehicle assignment and/or piece of work)</li> </ul>
Vehicle Monitoring	<p>Track vehicle location (and vehicle performance).</p> <p>Track vehicle performance relative to road type, pavement type, grade and elevation information.</p>	<ul style="list-style-type: none"> <li>• Street centerline (optional with pavement type, grade, elevation and hypsography)</li> <li>• Current location/time of transit vehicle</li> </ul>
Incident Management	<p>Locate vehicle. Translate latitude and longitude to landmark, intersection, or address (or visa versa).</p> <p>Route emergency or supervisor vehicle to revenue vehicle.</p> <p>Manage (transit) emergency vehicles at incident.</p>	<ul style="list-style-type: none"> <li>• Street centerline with addresses and landmarks</li> <li>• Incident location</li> <li>• Current location and status of emergency fleet</li> </ul>
Special Events Management	<p>Generate detour or alternative trip(s) to support planned event such as sports event or parade. Add bus stops, additional revenue vehicles, and route patterns.</p> <p>Generate maps and publicize alternative routes.</p>	<ul style="list-style-type: none"> <li>• Street centerline</li> <li>• Bus stop inventory</li> <li>• Route and route patterns</li> <li>• Time point inventory</li> <li>• “Patch” or schedule detour</li> <li>• Rescheduled ordered sequence of bus stops</li> <li>• Rescheduled ordered sequence of time points (and their trip time point times)</li> </ul>

<b>Supporting Task</b>	<b>GIS Application/Function</b>	<b>Transit Feature(s)</b>
Dynamic Re-routing and Timed Transfer of Trips, Vehicles, Blocks and Routes	<p>Automate the generation of detours.</p> <p>Generate new trip times based on traffic congestion, road geometries, number of lanes, road grade.</p> <p>Identify temporary stop locations based on population served, access, pad type, curb type, mode transfer (e.g., to commuter rail)</p>	<ul style="list-style-type: none"> <li>• Street centerline</li> <li>• Intersection diagrams</li> <li>• Photologs</li> <li>• Bus stop inventory (areas and transfer points)</li> <li>• Rail alignment (centerline)</li> </ul>
Transit Signal Priority	<p>Determine need for priority treatment (based on schedule adherence, load, etc.)</p> <p>Detect priority location</p> <p>Determine rolling average speed/interval for “queue jumping”<sup>2</sup></p>	<ul style="list-style-type: none"> <li>• Street centerline</li> <li>• Time Service Desired and Time Estimated Departure<sup>3</sup> <ul style="list-style-type: none"> <li>– Intersection (signal) stop bar</li> <li>– Location of end of queue</li> <li>– Need for priority treatment criteria (schedule adherence; load, etc.)</li> <li>– Traffic impedance (on links overlaid on route patterns)</li> </ul> </li> <li>• Rolling average interval (segment) and rolling average time to transverse segment</li> </ul>
Automatic Passenger Counting	<p>Aggregate boarding and alighting information at bus stop location and other locations.</p> <p>Provide load information between stops (intervals along route/pattern).</p> <p>Identify peak load location.</p>	<ul style="list-style-type: none"> <li>• Street centerline (with rail crossing and other locations vehicles are required to stop and open their doors)</li> <li>• Bus stop location/area</li> <li>• Route pattern</li> <li>• Trip (vehicle assignment or operator piece of work)</li> </ul>
Road Call Management	<p>Display distribution of road calls versus location, road calls versus routes;</p> <p>Generate route for road call service vehicles from origin to road call destination</p>	<ul style="list-style-type: none"> <li>• Street centerline with street addresses and landmarks;</li> <li>• Transit vehicle route, vehicle assignment, operator piece of work</li> </ul>

<sup>2</sup> Strategy for determining need for priority treatment. Data elements found in NTCIP 1406 TCIP On-Board Objects; strategy discussed in TCIP2 Dialogs Cluster #6 White Paper on Transit Signal Priority, version 2.4 (April 4, 2001).

<sup>3</sup> Based on NTCIP 1211 Signal Control and Prioritization (SCP) draft standard requirement.

<b>Supporting Task</b>	<b>GIS Application/Function</b>	<b>Transit Feature(s)</b>
Traffic Management	<p>Integrate and overlay traffic and road conditions with blocks, trip times (for schedule adherence, detours, etc.)</p> <p>Calculate average running time over link; calculate rolling average traversal of traffic management segment (as probe data).</p>	<ul style="list-style-type: none"> <li>• Street centerline</li> <li>• Transit vehicle route, vehicle assignment, operator piece of work</li> <li>• Current road and traffic conditions</li> </ul>
Grade Crossing Management	<p>Display vehicle location relative to grade crossing.</p> <p>Estimate time of arrival at crossing.</p>	<ul style="list-style-type: none"> <li>• Street centerline</li> <li>• Railroad crossings</li> <li>• Railroad alignments with signal system</li> <li>• Current rail vehicle speed; current transit vehicle speed</li> </ul>

**Table G2. Automated Traveler Information System**

<b>Supporting Task</b>	<b>GIS Application/Function</b>
Trip Itinerary Planning	<p>Generate route alternative trips based on spatial factors such as shortest walking distance, shortest trip distance, shortest (time or distance) trip, fewest transfers, obstacles to bus stops, nearest bus stop location, and landmarks.</p> <p>Generate return trip plan</p>
Schedule and Fare Information	<p>Display and query bus stop information.</p> <p>Estimate arrival/departure time of transit vehicle at bus stops not coincident with time points.</p> <p>Develop route maps.</p> <p>Quick development of service change maps.</p>
Operating Status	<p>Calculate time of arrival based on real-time running times, vehicle location, and distance between origin and destination.</p> <p>Provide displays of street closures and reroutes.</p>
Customer Comments	<p>Geocode customer address and/or employer location.</p> <p>Identify location referred to in the comment (if necessary).</p>
Kiosk Location Planning and Management	<p>Locate best location for information kiosk based on ridership and trip planning Origin/Destination analysis.</p>
Annunciator and Sign Management (on-board trip information)	<p>Trigger announcement and sign information based on vehicle location</p>

**Table G3. Electronic Fare Payment**

<b>Supporting Task</b>	<b>GIS Application/Function</b>
Fare Structure Flexibility	Accurately calculate distance and zone based on fare tables.

**Table G4. Transportation Demand Management**

<b>Supporting Task</b>	<b>GIS Application/Function</b>
Demand Responsive Reservation and Dispatch	Locate and display customer origin-destination.  Efficiently route Demand Responsive vehicle between multiple origins and destinations.  Generate manifest directions and hard copy map for driver.  Display distribution of DR trips.  Determine eligibility of ride (e.g., quarter mile from bus stop).
Ridesharing	Match rideshare partners (determine proximity of Origin/Destination pairs for dynamic and pre-determined rideshare partners).
Coordination	Match trips between Paratransit zone service providers.  Identify the best pickup and drop-off points for service provider coordination.

### **By Transit Business Function**

The following tables A-2 through A-13 present a broad set of examples of spatial data based transit applications or functions organized by transit business area.

**Table G5. Service and Capital Planning**

<b>Supporting Task</b>	<b>GIS Application/Function</b>
<i>Service and Route Facilities Planning</i>	
Produce service change process materials.	Produce maps of selected service areas and display route proposals.
Presentations to Council, community groups.	High quality maps displaying streets, facilities, proposed routings, current and deleted route portions.
Prepare analyses of the service operating environment for use in planning changes to transit service.	<p>Display current service route patterns and facilities including park-and-ride layouts, transit center layouts, layover zones, etc.</p> <p>Display related service statistics such as vehicle volumes by time of day or number of routes using a street.</p> <p>Display current or future infrastructure such as street class, street width, street direction (if one-way), operating speed, parking conditions, pedestrian barriers, grades (slopes), HOV lanes, turning radius, curb cuts, signals, level of service, etc.</p> <p>Display current and past ridership by route, route portion, bus stop or other geographic area.</p> <p>Highlight routes or route portions based on productivity levels.</p> <p>Display information related to ridership demand such as population, employment, travel patterns, population and employment densities, land use designations.</p> <p>Plot and analyze address lists of major employers and institutions.</p> <p>Combine and display the above listed information using digital maps, tabular lists, video, photos, documents.</p> <p>Provide demand modeling of transit service,</p>
Respond to overloads.	Display ridership overloads by route or route portion, using different colors, line widths, symbols, and assorted graphics. Combine these with scheduling information to display headways and vehicle types.

<b>Supporting Task</b>	<b>GIS Application/Function</b>
Plan zones and shelters, including adding, deleting and moving facilities.	<p>Update the base map of zones and shelters. Display new routes with existing bus zones or shelters.</p> <p>Display related information on bus zones and shelters such as property ownership (tie-in to assessor's data), site layout, passenger boarding and alightings, etc.</p>
Maintaining bus zone history.	Use multi-media to maintain zone history including changes, legal documents, site plans and photos.
Analyze bus zone spacing.	<p>Display bus zones with passenger usage.</p> <p>Highlight bus zones based on spacing from adjacent zones. Combine spacing information with passenger usage information.</p>
Transit Flow and Safety Program (TFASP) planning.	<p>Use vehicle volumes to flag high utilization locations.</p> <p>Analyze and map elements of accident sites to improve zone site selection, distance from turning points.</p> <p>Use multi-media to tie drawings, aerial photos, etc. to the location.</p>
Leased lot planning (supplemental park and ride lots).	<p>Update the base map of leased lots. Display new routes with existing leased lots.</p> <p>Display related information on leased lots such as property ownership (tie-in to assessors data), site layout, etc.</p>
<i>Park-and-Ride Program</i>	
Determine optimum site, size, configuration, etc. for park-and-ride lots.	<p>Map current and potential park-and-ride sites, capacity, usage and other associated data such as ownership, legal address, survey data, etc.</p> <p>Display park-and-ride data against service patterns, volumes, and passenger volumes.</p> <p>Estimate expansion or new site potential using demographic information and creating customer draw areas (buffering).</p> <p>Map land use (zoning) and available land including transit owned property.</p> <p>Tie-in to tax assessor's data to assist in property owner contacts.</p> <p>Geocode addresses based on license plate surveys to determine park-and-ride draw area.</p>
<i>Passenger Shelters</i>	
Determine optimum site, size, configuration, etc. for shelters.	<p>Display current shelters and zones, with associated usage.</p> <p>Tie-in to assessor's data to assist in property owner contacts.</p>



<b>Supporting Task</b>	<b>GIS Application/Function</b>
<i>Urban Arterial Improvements</i>	
Identify infrastructure projects with potential benefits for public transit.	<p>Display adjacent land use and demographics, transit route patterns, vehicle volumes and existing road types.</p> <p>Displays of travel time and Level of Service along certain street segments to identify key bottleneck locations.</p> <p>Tie-in to aerial photos, schematic representations of target arterial and surrounding land for preliminary planning.</p>
<i>Long Range Planning</i>	
Service and facilities planning.	Applications similar to short-range planning.
Develop service plans that support the recommended rapid transit system investment.	<p>Display geographic subareas for resource or subsidy investment. For each subarea and jurisdictions within subareas, compare current and future subsidies against current and future equity factors such as farebox revenue, sales taxes, population, employment, household densities, low income population, etc. Create different subareas as needed.</p> <p>Calculate coverage of service against population (e.g., show what percentage of population in a subarea is within 1/4 mile of a bus route, calculate how much of a route's length is within a certain subarea).</p>
<i>Long Range Forecasting</i>	
Presentation maps and graphics for long range forecasting models.	Provide the base street and service network for future subsidy and revenue allocation models.
Forecast ridership, productivity of various implementation options via network modeling (EMME/2 or other package).	Take downloaded information from forecasting package and create displays of trip generation and travel pattern data by TAZs or other geographic boundary areas. Overlay and combine with other data such as demographics or transit route patterns and facilities.

**Table G6. Scheduling**

<b>Supporting Task</b>	<b>GIS Application/Function</b>
General route scheduling	<p>Provide street or route networks needed by new runcutters or scheduling databases.</p> <p>Display service routes, deadhead routes, street network.</p> <p>Display running times and variability of running times by time of day. Display headways by route or street segments.</p>
Estimated time point development	Display bus zone or route level ridership to aid in estimated time point development for outbound service.
Turnback analysis	Display passenger loads by bus zone or route segment to gauge impact of proposed turnbacks.
Through-routing analysis	Display origin-destination data to measure customer impacts of route splitting.

<b>Supporting Task</b>	<b>GIS Application/Function</b>
Dead-head routing	Use network optimizers to determine shortest path.  Display infrastructure attributes to show restrictions on routing.
Base assignment (best base to run service from)	Provide street and routing networks needed for potential optimization models.
Road relief determination	Provide street and routing networks needed for potential optimization models.
Running time analysis	Compare bus running times on street segments that overlap multiple TPIs to check for consistency.  Display on-time performance at selected locations.  Display changes in running time from past signups.  Compare running time changes with traffic volume changes.
Passenger count reporting and analysis	Display passenger counts by route, checkpoint, cordon line, screenline or screenline defined area (e.g. CBD).  Display overload routes and route segments by coach type. Display excess seating capacity at selected locations.  Display route peak load points.

**Table G7. Market Development**

<b>Supporting Task</b>	<b>GIS Application/Function</b>
Develop transportation management plans with governmental agencies, private developers and private sector employers.	Display and analyze existing and planned infrastructure and transportation environment. Similar to service planning needs (e.g. streets, transit routes, transit facilities, existing mode splits, etc.).  Display major employer sites. Tie to employer attribute files containing information such as type of business, parking conditions, mode split, and number of employees.  Display employer sites against available transportation services.  Display various land use conditions (e.g. parking supply and cost, permits, zoning).
Assist employers in meeting Commute Trip Reduction Law (if applicable).	Create targeted market mailing lists, either of employers or employees based on geocoding analyses.  Create and display Commute Trip Reduction (CTR) zones (based on traffic analysis zones) with average vehicle miles traveled (YMT) and SOVs.

Supporting Task	GIS Application/Function
	Geocode employee address information.
	Create customized maps for individual employment sites.

**Table G8. Accessible Services**

Supporting Task	GIS Application/Function
<i>Supplemental or Paratransit Services Program.</i>	
Develop and operate program in response to the requirements of the Americans with Disabilities Act (ADA).	
Determine applicant eligibility.	<p>Process trip origin-destination pairs against existing service to determine if a trip falls within the 3/4 mile eligibility boundary (or provide street and route network to other system).</p> <p>Geocode ADA applicant addresses and perform location related analysis.</p>
Maintain or provide support for trip planning software.	Provide access to the current base map of streets, the regular route network, and geocoded address ranges. This is especially important to keep trip planning software current with service changes (route and bus stop) that affect eligibility areas.
Program evaluation.	Calculate spatial quantifiers such as total number of passenger miles produced, average distance per trip, passenger volumes to/from specified boundary areas, etc.
Program planning, provide/support transportation broker tools.	<p>Display and analyze existing and planned infrastructure, transportation services, facilities, demographics, land use factors and forecasted population growth by class (e.g. elderly, low income etc.). Similar to service planning needs.</p> <p>Display topographical data as well as other potentially ADA related information such as curb cuts, stairways, barriers, etc.</p>
Track other accessible services.	Display and analyze lift usage, lift problems on regular service.

**Table G9. Vehicle and Facilities Maintenance**

<b>Supporting Task</b>	<b>GIS Application/Function</b>
<i>Vehicle or Coach Maintenance.</i>	
Track vehicle mileage for preventative maintenance schedules.	Track mileage by vehicle and route.
Track and analyze mileage between trouble calls.	Geographically analyze trouble calls, by location, type of vehicle and type of trouble call to determine if any patterns exist.
Locate vehicles in a yard or facility.	Map and track vehicles within a yard or facility.
Analyze lift malfunction locations to determine if any location-related problems exist.	Create maps to analyze lift malfunction locations.
Tow vehicles back to base maintenance.	Geographically analyze wrecker dispatch calls to determine optimum location from which to dispatch crews. GIS-based shortest path models will be used to minimize distance traveled and response time.
Reduce vandalism on coaches.	Match vandalism incidents with coach assignment lists and display the most troublesome locations and routes.
Estimate future mileage on coaches.	Use the GIS database to calculate the mileage increases or decreases resulting from route change proposals.
<i>Trolley Overhead (TOR) and Electrical Maintenance.</i>	
Respond to TOR work orders.	Create a work order packet containing an area map to assist work crews in locating the correct poles, wires, etc. Include operation and maintenance documents related to the work order site, retrieved via multi-media features.
Repair TOR	Compare trolley vehicle volumes with work order information to determine relationships between wire usage and wear.
Forecast work load.	Use current and planned trolley service levels and TOR expansion estimates to forecast the amount of TOR inspection and repair work required, and the associated staffing needs.
<i>Building and Facilities Maintenance.</i>	
Plan shelter cleaning.	<p>Use GIS map-based network optimizing module to determine shortest path covering all designated shelters.</p> <p>Perform "what if" analyses to model the effects of different cleaning cycles, more cleaning crews or additional shelter installations.</p> <p>Provide large maps of service area with routes and facilities.</p>
Provide feedback to planning on facilities vandalism.	Show frequency and type of vandalism by location.
Maintain buildings and grounds.	Use GIS multimedia capabilities to store and retrieve floor plans, landscape plans, and plans for irrigation systems, storm water systems, fire systems, shutoff valves, etc.

<b>Supporting Task</b>	<b>GIS Application/Function</b>
<i>Engineering and Facilities Support.</i>	
Planning for excavation work.	<p>Tie-in to utilities databases from other agencies to see where excavation work will require utilities movement.</p> <p>Tie-in to construction as-built engineering drawings and Engineering AutoCAD files.</p>
Coordinate ongoing construction among other agencies.	Tie-in to GIS databases from other agencies and companies to anticipate future construction plans that may affect Metro service and property.
Other facilities maintenance tasks.	Additionally, GIS may have application to park-and-ride lot maintenance, hazardous waste removal, and tracking vandalism.

**Table G10. Sales and Customer Services**

<b>Supporting Task</b>	<b>GIS Application/Function</b>
Advertising and promotions.	<p>Use geocoding to create customer mailing lists for route specific marketing.</p> <p>Designate heavy use stops in target areas for promotions.</p> <p>Use GIS spatial functions to identify and quantify the draw area of various park-and-rides. Use identified area for target marketing.</p>
<p>Sales</p> <ul style="list-style-type: none"> <li>Includes management of retail pass sales outlets</li> </ul>	<p>Use GIS as a foundation for future models which might create customer profiles (pass users, cash riders, etc.) by combining demographic information with passenger usage.</p> <p>Create system maps and special area maps tailored to different customer groups.</p> <p>Map geographic distribution of outlets.</p> <p>Display outlet performance statistics by location, region or outlet type.</p>
Customer Services	See Table 3-1d, Automated Traveler Information System. Provide tools for both customers and Rider Information Office staff.
Maintain schedule holder information and timetable racks.	<p>Maintain inventory of bus stop signage (pole type, schedule or information holder type, sign type, etc.) and provide crews with maps and shortest path info.</p> <p>Display timetable usage for distribution outlets. Tie to inventory system to display which timetables are replaced most and what timetables are not used.</p>
Ridematching.	See Table 3-1d, Transportation Demand Management.

**Table G11. Vanpool Program**

<b>Supporting Task</b>	<b>GIS Application/Function</b>
Vanpool formation planning.	Analyze ridematch and vanpool applicant locations to determine optimum collection points.
Increase vanpool flexibility	<p>Display vanpool route patterns and unmatched applicants to identify potential mid-route pickup riders.</p> <p>Provide van groups needing more riders with maps showing other potential riders from ridematch system. Potential riders could be screened by distance criteria.</p> <p>Display employer sites with appropriate information (number of employees, business type, business hours, parking conditions, etc.).</p>
Assess the ridematch and vanpool potential at employer sites and promote vanpools.	Geocode employee addresses and flag potential matches.
Vanpool operations and maintenance.	Use GIS spatial analysis tools to locate optimum garage locations for each van.
Vanpool safety.	Display and analyze vanpool accidents for any trends.

**Table G12. Research and Evaluation**

<b>Supporting Task</b>	<b>GIS Application/Function</b>
Provide information from on-board Origin and Destination surveys.	<p>Geocode respondents' origins and destinations to create travel pattern data.</p> <p>Display results visually by route. (e.g. route colors, line patterns indicate number of riders, average fare payment, modal age, etc.).</p>
Assist with the design, implementation and analysis of route surveys.	Access route level information such as vehicle volumes and passenger counts for sample sizing and selection.
Evaluate effectiveness of various service and facility improvements.	Compare changes in ridership, travel patterns and other data.
Conduct studies for transportation projects at major institutions and employers	<p>Display target population/market travel patterns and home locations using geocoding.</p> <p>Overlay employee travel pattern needs with available transportation services (routes, vanpool groups, etc.).</p> <p>Overlay other information as needed such as demographics and land use.</p>
Obtain customer and non-customer feedback through survey research.	<p>Compare sample against population for sample areas to determine sample representation quality.</p> <p>Produce maps displaying survey results for reports and presentations.</p>
Show environmental scan results for strategic and management planning, including analyses of potential markets.	Provide maps displaying pertinent information such as demographics, employment dispersion, service levels, activity centers, land use, etc.

<b>Supporting Task</b>	<b>GIS Application/Function</b>
Display and analyze data collected from registering fareboxes.	Display route maps showing boardings and fares for each route.
Complete the Title VI Equity in Transit report	Analyze and display regional demographic data.

**Table G13. Transit Operations**

<b>Supporting Task</b>	<b>GIS Application/Function</b>
<i>Bus Operations</i>	
Help automate production of driving instructions.	Create quick information maps/guides for operators to help answer customer questions.
Provide operators with route information	Create large service maps for posting in window area showing routes, reroutes, snow routes, construction, etc.  Create maps for The Book (operators' instruction/information book).
Locate vehicles in a yard or facility.	Map and track vehicles within a yard or facility.
Relief point and relief car planning.	Provide street, route and distance information.
Work force estimation	Perhaps GIS could provide the street and route network needed for work force models and runcut models. The models will be used to estimate service change proposal costs.
<i>Operator Training</i>	
Qualify operators on routes.	Provide timely and accurate maps for trainers and operators.  Use multimedia with GIS to allow operators to "drive" the route via a computer display, see potential trouble stops and allow for queries.
<i>Subcontracted Service</i>	
Service planning.	Applications similar to directly operated service.
Contract performance monitoring.	Display results of supervisor field checks and customer complaints by location and contractor.
Demand response service analysis.	Display and analyze dial-a-ride origin and destination information.
<i>Street Supervision</i>	
Optimize the size of street supervisor districts.	Display and analyze supervisor work records by location.  Create color-coded maps showing supervisor areas by day and time (e.g., night supervision areas, weekend coverages etc.).
Analyze supervisor activity data.	Geographically display and analyze supervisor activity information, such as lift-assistance calls, security calls, other mechanical related calls, and special assistance trips (covering for missed service).

<b>Supporting Task</b>	<b>GIS Application/Function</b>
Plan for snow operations.	<p>Display and analyze prior snow related problem locations. Include in the history the type of equipment.</p> <p>Display and analyze route coverages to determine which routes might be deleted during snow while maintaining good service coverage.</p> <p>Provide maps showing where sand is stored.</p> <p>Show street and route system with topography to assists snow route planning.</p>
Comfort station planning.	Display existing comfort stations by time restrictions and compare against routes.
<i>Radio and Other Communication</i>	
Communications coverage.	Develop communications infrastructure coverage areas based on topographical maps to identify target areas not covered by the existing and planned communications towers or commercial providers.
Operate radio system Automatic Vehicle Location (AVL) system.	Provide base street and route map network for AVL. May include on-line real-time map and location displays.
Monitor service performance.	<p>Display locations, routes and on-time performance statistics.</p> <p>Display calls related to delays (security, mechanical, accidents, other operating). Maybe combine with traffic flow data from state or other jurisdictions.</p>
Respond to operator needs in field.	<p>Display service supervisor districts with routes.</p> <p>Display fire district or E911 coverage areas.</p> <p>Display detailed street map showing where reroutes might be taken.</p> <p>Use shortest path network modules to assist in reroutes.</p> <p>Show trolley overhead system and switches.</p>
Coordinator workload balancing.	Analyze calls by area to balance number of coordinators per channel.



**Table G14. Risk, Safety and Security**

<b>Supporting Task</b>	<b>GIS Application/Function</b>
<i>Risk Management, Insurance and Safety</i>	
Analyze accident data.	<p>Enter accident location directly to a map without requiring extensive address creation.</p> <p>Display accident locations according to various attributes (time of day, type of accident, type of vehicle, etc.).</p> <p>Display highest accident locations by various attributes.</p> <p>Use GIS-related multi-media systems to tie-in engineering drawings or aerial photography of street layout near an accident site to examine corner geometry , lane widths, etc.</p>
Create performance statistics	Calculate indicators such as accidents per million miles within specific geographic areas or at specific geographic locations.
Support disaster planning.	Display topography, critical structures, national geodetic survey data, emergency shelters, emergency personnel contacts, etc.
Safety classes and presentations.	Create high quality maps displaying accident information for various meetings.
<i>Security</i>	
Plan work force requirements and security personnel assignments (facilities, boarding teams, coverage areas).	<p>Display security incident locations with associated attributes (time, type of incident, injuries, property loss, number involved, etc.). Especially valuable for examining historic data to develop trends.</p> <p>Display and analyze demographic data, especially forecasted changes in population, employment, income, etc.</p> <p>Tie-in to other security force databases (sheriff, police, etc.) to develop a better overall picture of security problems.</p>
Educate operators on where problems may occur.	Create large maps for operator training sessions or displays.
Future real-time location of transit security officers.	Provide base map necessary <i>for</i> global positioning system technology.
Manage security data	Use GIS-related multi-media systems and database capabilities to tie security incident locations to associated pictures, documents, reports, information. (Note: Security data is highly confidential and GIS should provide safeguards against unauthorized use.)

**Table G15. Communications and Community Relations**

<b>Supporting Task</b>	<b>GIS Application/Function</b>
<i>Public Participation</i>	
Contact citizens, community groups and other public groups regarding projects and service changes	Create mailing lists or identify groups to be contacted by overlaying geographic buffers on target areas. (For example, identify addresses within 1/4 mile of a route change or create a one mile radius ring around a proposed facility and identify all community and neighborhood groups which are touched by the ring.)
Respond to customer complaints and questions from board members and elected officials.	"Geocode" complaints (tag with a geographic location code based on the address) <i>for</i> cataloging and analysis.  Summarize and display complaint information by various boundary categories (e.g. political districts, incorporated areas, Forecast Analysis Zones, etc.).  Display information on operating environment that is relevant to the question or complaint. Similar to route planning GIS application.
<i>Staff Information and Communication</i>	
Create graphics as needed for documents, public meetings, etc.	Use GIS hardware, software and database to create maps and documents as needed.

**Table G14. Property and Right-Of-Way**

<b>Supporting Task</b>	<b>GIS Application/Function</b>
Buy property.	Display vacant land and properties listed <i>for</i> sale (subject to data availability).  Use spatial analysis tools and tie-in to assessor's database to estimate land values, zoning codes and registered owners for future bases, park-and-ride lots, stations and rapid transit corridors.  Combine ability to display assessors and real estate sellers data with overlays of proposed Transit construction facilities to determine potential for property acquisition within a certain distance of proposed sites.  Also overlay other information such as wetlands, hazardous waste or sensitive area classifications.
Lease property.	Display properties available for rent or purchase (subject to data availability).
Sell property.	Maintain map and information pertaining to transit owned property.
Provide relocation assistance to displaced owners and residents.	Tie-in to real estate sellers databases for potential properties to relocate persons, businesses. Overlay with other land use information as specified by displaced parties such as schools, transit access, parks, etc.

## **APPENDIX H    CONCEPT OF OPERATIONS FOR TRIP ITINERARY PLANNING**

### **Trip Itinerary Planning Use Case**

[from *Geospatial One-Stop Transit-Model Advisory Team*]

#### **Overview/Description**

A customer seeks information related to point-to-point travel plans that includes public transportation services such as bus, rail or other mode. The trip itinerary request may be based on several key criteria such as origin, destination, travel date/time, amenities desired, traveler profile, trip constraints (i.e. lowest cost, shortest time, fewest transfers, mode, accessibility, time of day, day of week), and one way or return trip. Trip plan may include real-time information on schedule adherence, route adherence and service changes (see Rerouting Use Case) due to “incidents.”

This use case does not include paratransit or demand responsive scheduling and dispatch requirements.

#### **I. Concept of Operations**

1. Customer generates a trip request identifying origin, destination and time/date of travel (potentially specifying preference criteria pertaining to cost, transfer, mode/carrier, etc.).
2. The Trip Planning System verifies that request is complete and accurate. When verified it processes customer request and generates a trip plan.
3. The Trip Planning System verifies that the scheduled data is still valid on all legs for near term (review trip plan against reroutes, updates and planned events). If there are exceptions, the system regenerates the itinerary.
4. For trip plans in the near future, the system checks real-time status of service. If there are exceptions, it regenerates the itinerary.

#### **Enumeration of Needs**

The following functions are needed to respond to customer requests for trip itinerary requests.

#### **Request and verify customer trip criteria**

- Provide options for selecting origin and destination
  - Potential origin and destination names should be comprehensive including addresses, vanity addresses, intersections, landmarks (e.g., malls, squares, hospitals, etc.), community centers (e.g., Hyde Park in Chicago).
- Provide service and amenity options for different modes
  - Date

- Departure or Arrival time
- Service area
- Routes/trips
- Public Transportation Stops
- Provide options for criteria selection (“minimizing”) including
  - Shortest trip (travel time , trip distance)
  - Shortest walking/driving distance (access/egress)
  - Least number of transfers
  - Least costly
- Provide options for including/excluding various criteria including
  - Mode
  - Via Landmark
  - Amenity at stop point (e.g., parking availability, accessibility)
  - Service types (e.g., express or local)
  - Route
  - Fare Media (e.g., pass, cash, credit card)
  - Accessibility

**Provide transit trip itinerary plan based on customer criteria**

- Provide driving or walking/biking directions and distances to/from selected public transportation stops to/from origin/destination
  - Identify obstacles, barriers, accessibility, amenities for directions to/from public transportation stops
- Provide directions for walking between transfer points
- Provide transit information on planned bus route numbers, travel direction, schedules and current operations including route numbers, departure/arrival times and locations, and transfers within a mode and between modes
- Provide service information on different modes including planned and unplanned detours and real-time schedule adherence information
- Provide amenity information on public transportation stops
- Provide fare cost information for planned trip
- Provide estimated travel time for the itinerary
- Provide a written trip itinerary summary containing the sequential unlinked trip making activities including origin, boarding bus stop location and ID number, first transit route, alighting location, transfer to 2<sup>nd</sup> bus route information including boarding and alighting, and so on, until the final destination.

To meet these user needs, the following data needs should be supported:

- Provide a Topologically Complete and Logically Consistent Transportation Network including street names and addresses, alternate street names, cities, zip codes, barriers
- Provide a Complete List of Street Names and Landmarks including park and rides, transit centers, and neighborhood locations
- Provide a Complete and Logically Consistent Transit Network and Features
  - Patterns, public transportation stops and time points over all transit modes, and transfer points

- Revenue trip times (particularly detailed bus trip schedules referenced to trip pattern spatially on a stop-by-stop or time point-by-time point level), and estimated transfer and wait times at stop points (by time of day)
- Provide Fare Information for all combinations of itineraries (including transfers)
- Provide Real-time schedule adherence data
- Provide park-n-ride , transit center data on location, size, amenity, and other characteristics description
- Provide bus stops list on ADA, amenity, shelter, and other relevant characteristics.
- Provide Unplanned rerouting information (see Unplanned Re-routing use case)

## **II. Functional requirements for supported operation**

### **Overview of requirements**

The requirements for supporting the Trip Itinerary Plan Use Case are

- Request and verify trip itinerary request
- Provide trip itinerary plan to customer

### **Detailed functional requirements**

#### **Trip Itinerary Planning (TIP) Functional Requirements**

<b>Requirements</b>
<b><i>TIP.1—Request and verify trip itinerary criteria</i></b>
<p>TIP 1.1—The system shall provide options to the customer to create an itinerary request. The information components needed for this requirement are:</p> <ul style="list-style-type: none"> <li>▪ Information on service area, modes, routes, public transportation stop points</li> <li>▪ Information on origin, destination, date and time of travel</li> <li>▪ Information on service types and trips per route (including attributes of each trip in the route, e.g., wheel chair accessibility, bike or ski rack)</li> <li>▪ Information on fare media accepted</li> <li>▪ Information on public transportation stop points, their amenities and accessibility</li> <li>▪ Information on allowable criteria selection features (include and exclude; minimize)</li> </ul>
<p>TIP 1.2—The system shall verify that the customer request is complete and accurate. In addition to the information components listed in TIP 1.1, this requirement needs the following information components:</p> <ul style="list-style-type: none"> <li>▪ Information on addresses, vanity addresses, landmarks, alternate street names <ul style="list-style-type: none"> <li>○ Complete List of Street Names and Landmarks</li> </ul> </li> </ul>

**TIP.2—Provide transit trip itinerary plan to the customer**

TIP 2.1—The system shall develop a transit trip itinerary based on customer criteria. To accomplish this, the systems must determine several alternative itineraries. The internal functions are

- The location of entry and exit into the transit network
- The most efficient path from entry to exit in the transit network (based on selected criteria), this may require incorporating walking and wait times at transfer locations
- Walking directions for transfer
- Walking and driving directions from/to origin/destination to/from entry/exit points including barriers, obstacles and modal network connections within the *transportation* network
- List of amenities at specific public transportation stop points associated with plan
- Calculate fares and allowable fare media (including transfers) per leg as well as total cost
- Calculate total travel time
- Calculate total walking time
- Generate return trip (if requested)

The information components needed for this requirement are:

- For Origin/Destination and Walking and Driving Directions –
  - Topologically Complete and Logically Consistent Transportation Network including street names and complete addresses
- For Developing the Transit Trip Plan
  - Complete and Logically Consistent Transit Network and Features
    - Patterns, public transportation stop points and time points over all transit modes, and transfer points
    - Revenue trip times (each bus trip can be geo-referenced by trip pattern spatially on a stop-by-stop or time point-by-time point level), and estimated transfer and wait times at stop points (by time of day)
  - Fare Information for all combinations of itineraries (including transfers)

TIP 2.2—The system shall verify the trip itinerary plan against any planned or unplanned detours, delays or special services. The information components needed for this function are:

- Unplanned rerouting along the trip plan
- Status of Planned/construction of RoadSeg along the trip plan
- Special service schedules (auxiliary parking facilities that are used as pick-up/drop-off points)

The information components needed to support this function are not included in the scope of the Data Mapping section.

TIP 2.3—The system shall provide real-time updates on itineraries that are scheduled for the near future (e.g., within an hour). The information components needed for this function are:

- Estimated departure/arrival times of vehicles designated to perform selected trips in itinerary.

The information components needed to support this function are not included in the scope of the Data Mapping section.

### ***III. Mapping data requirements to current transit model***

The data requirements necessary for trip itinerary planning use case are described in the table below.

Data Requirements	Definitions	Assumptions/Conditions
Public Transportation Stop	An established location where public transport customers may board or alight from a transit vehicle in revenue service.	Attributes: Bus Stop Inventory attributes including: <ul style="list-style-type: none"> <li>▪ Wait times</li> <li>▪ Fare Zone (if applicable)</li> </ul>
Amenity	The elements of a physical feature, a fixed location, or a transit facility. The amenities of a public transportation stop, for example, may include the shelter, platform announcement panel, and benches. An amenity may be described by one or more characteristics, or attributes, such as the year of construction or its current condition.	Unique ID, Name, type. Domain types: list the known types and add 'other'.
Transfer points	A transfer point is a geographic location that encompasses more than one bus stop where a customer can move from one route to another. Defined the same as cluster.	A transfer point could have an associated impedance, if so, it might not be accessible. Transfer points or clusters need to be composed of PTStops that make up a transfer cluster, and unique identifier. Impedance, walking directions between stop points should also be included

Data Requirements	Definitions	Assumptions/Conditions
Patterns (including routes)	A unique, non-branching, ordered sequence of time points, street links, public transportation stops from the beginning of a route to the end of a route.	<p>Patterns provide the path of the trips, temporal service information is contained in the trips. Patterns serve a function similar to ‘anchor points’ in the Geospatial One-Stop Road Model Advisory Team standard. The density of time points is directly related to the accuracy needed in the schedule.</p> <p>The pattern is an ordered sequence of road segments, as well as time points, and stops. But transit agencies usually use the whole segment, rather than portions of segments, even when the PTStop occurs in the middle of the block.</p> <p>Fields include:</p> <ul style="list-style-type: none"> <li>▪ ID</li> <li>▪ Route number ID</li> <li>▪ Service type – local/express</li> <li>▪ Schedule version</li> <li>▪ Revenue/non-revenue</li> </ul> <p>Model Element:</p> <ul style="list-style-type: none"> <li>▪ Ordered sequence of streets (Road_seg) that make up the pattern, which furnishes geometry</li> </ul> <p>Note: At WMATA path/route is an ordered sequence of time points. For trip planning there is an ordered set of road segments or an ordered set of time points.</p>
Streets		<p><b>Street links:</b></p> <ul style="list-style-type: none"> <li>▪ ID</li> </ul> <p>Road_Seg:</p> <ul style="list-style-type: none"> <li>▪ Status</li> <li>▪ authorityID</li> <li>▪ fieldMeasure</li> <li>▪ length</li> <li>▪ geometry (optional)</li> <li>▪ topology (optional)</li> <li>▪ isAnchorSection (Boolean)</li> </ul> <p>All other attributes that pertain to roads are classified as linear or point events.</p>
Address Ranges		See Address Addendum



Data Requirements	Definitions	Assumptions/Conditions
Landmarks		<b>Landmark attributes:</b> <ul style="list-style-type: none"> <li>▪ Name</li> <li>▪ Type</li> <li>▪ Location</li> </ul>
Parking lots		Attributes: <ul style="list-style-type: none"> <li>▪ ID</li> <li>▪ Public Transportation Stop(s) served</li> <li>▪ Transit routes served and schedules</li> <li>▪ Owner (optional)</li> <li>▪ Facility phone (optional)</li> <li>▪ Total Spaces</li> <li>▪ Operating Hours</li> <li>▪ Parking Provided for each vehicle class (spaces, rates, permissible entrances, fill time, other information) (optional)</li> <li>▪ Availability of charging facility for electric cars (optional)</li> <li>▪ Bicycle storage and lock facility (optional)</li> </ul>
Walking distance	Road_Seg	
Accessibility		Curb cuts and corners are needed for ADA compliance. They should be attached to the road segment.
Obstacles to walking, grade data		An attribute is added to the Road_Seg; “walking permitted – yes/no”. The attribute is applied to each individual road segment. An obstacle could be a linear event, but is often just a Boolean attribute on a segment. Typically, one can walk along or walk over. The value applies to the entire segment.
Fare data (based on distance, or zones, or flat)		Types of fare policies: <ul style="list-style-type: none"> <li>▪ Flat</li> <li>▪ Distance</li> <li>▪ Zone</li> <li>▪ Time of day</li> <li>▪ Transfers</li> <li>▪ Special rider classes (elderly, youth, disabled)</li> <li>▪ Bulk pass discount</li> </ul> Six-dimension table is required to handle the relationships between all the variations in fare types.  See Addendum for guidance on TCIP Fare Collection Fare Tables (NTCIP 1408)

Data Requirements	Definitions	Assumptions/Conditions
Schedules	From TCIP: “A table that includes all the time points and trips on a route. Contained within the SchRoute is the Master Schedule Header information. Contained within SchTrip is the day type information.”	For each route: <ul style="list-style-type: none"> <li>▪ Time Table Version</li> <li>▪ Activation date</li> <li>▪ Deactivation date</li> <li>▪ All supported trips assembled by route direction, service type, and day type</li> <li>▪ All supported trips in correspondence to the trip pattern number</li> </ul>
Trips	“A one way scheduled movement of a transit vehicle between starting and ending time points.”	Attributes: <ul style="list-style-type: none"> <li>▪ ID</li> <li>▪ Trip type</li> <li>▪ Ordered sequence of time points with their times (of arrival)</li> <li>▪ Pattern ID (associated with)</li> <li>▪ Time Table Version</li> </ul>
Pedestrian (sidewalks, bike paths, walking paths, centerline dividers)		Attributes: <ul style="list-style-type: none"> <li>▪ TBD</li> </ul>

## **APPENDIX I    TRI-MET'S CONFIGURATION MANAGEMENT PLAN FOR BASE MAP UPDATE**

This appendix describes Tri-Met's Configuration Management Plan. It supports update and distribution of base map features and attributes to all applications that use the base map. RASDA, PASS, and ATIS are names of applications supported by Tri-Met at the time of this document's publication. More information on developing a Configuration Management Plan is contained in Section 9.3.2.

### **Tri-Met Procedures to Prepare Base Map for RASDA, PASS, ATIS**

#### **ALL FILES**

1. **Project** Clark County Roads from Washington South to Oregon North.
2. **Merge** Streets coverage from Metro & Clark County Roads shapefile together maintaining attribute table from Metro.
3. **Populate** all null records in:
  - Streetname to "Unnamed"
  - Type to 1500
  - County to "CLAR" (should be only nulls in file)
  - City to Cityname through spatial query on City Boundary poly.
  - Zipcodes to Zipcode (nulls not expected)
4. **Check** Type 2000s  
May need to move to Type 1500 all Type 2000s with modify date starting 6/2001
5. **Modify** bridge geography to connect
6. **Add** missing street segments required for RASDA (in-house ArcInfo app to maintain route topology)
7. **Build** line topology
8. **Copy file:**
  - TO\_RASDA/ str1101
  - TO\_ATIS/NOV01/ str1101.shp (final file str1101LL83.TAB)
  - TO PASS/ NOV01/ streets.shp (final file streetsLL27.TAB)

#### **TO RASDA**

(Cannot contain edits for ATIS)

Regenerate and Recalibrate routes using new base map.

#### **TO ATIS**

1. Delete Street Classifications (@4200 street segments not needed for routing):
  - 1600 – alley (@1600 segs)
  - 2000 - unimproved road way passable by emergency vehicle (Ptld only for 911) (@30 segs)

2. Convert from Shapefile to MapInfo file format using SHAPELINK utility  
Specify current coord system as StatePlane 83 ft Oregon North
3. In MapInfo, FILE - SAVE-COPY-AS  
Specify new coord system as Lat/Lon (NAD 83 for Continental US)

## **TO PASS**

1. Do NOT delete any Street Classifications
2. Shorten the following street names
  - Martin Luther King Jr. to MLK
  - Martin Luther King Jr. Blvd-MAD to MLK Blvd
  - Martin Luther King Jr. Blvd-MOR to MLK Blvd
  - Martin Luther King Jr. Blvd-UNI to MLK Blvd
  - Martin Luther King Jr. Blvd-UNI to MLK Blvd
  - Beaverton Hillsdale to Beav-Hills
  - Beaverton Hillsdale-Canyon to Beav-Hills-Canyon
  - Beaverton-Hillsdale-Hwy 217 to Beav-Hills-Hwy 217
  - US Veterans Hospital to US Veterans Hosp
  - Also check alleys for ALL streets (max characters for alleys is 22)
3. Standardize references to highways:  
All street names with HWY should be listed with an FTYPE as HWY:
  - Hwy 211 to Hwy 211 Hwy
  - Hwy 213 to Hwy 213 Hwy
  - Hwy 35 to Hwy 35 Hwy
  - Hwy 99E to Hwy 99E Hwy
  - Hwy 47 to Hwy 47 Hwy
  - 212-224 Hwy to Hwy 212-224 Hwy
  - 211-224 Hwy to Hwy 211-224 Hwy
  - US Hwy 26 to Hwy 26 Hwy
4. Generate new ADA Boundary  
Buffer 3/4 mile @ all fixed routes & cookie cut out of tm\_fill.shp
5. Convert ADA.SHP & STREETS.SHP to MapInfo format using SHAPELINK utility  
Specify current coord system as StatePlane 83 ft Oregon North
6. In MapInfo, FILE - SAVE-COPY-AS  
Specify new coord system as Lat/Lon (**NAD 27** for Continental US)

Mapping layers required for PASS (updates generated as boundary changes are made):

1. Centerline Street Base from Metro
2. ADA Boundary generate in house
3. Tri-Met District Boundary from Metro
4. Urban Growth Boundary from Metro
5. Zipcode Boundaries from Metro

Periodically Geocode PASS Master Address Table in ArcView and populate new columns:  
Xcoord, Ycoord, Lat, Lon, UGB, TSB, ADA



## APPENDIX J ACRONYMS AND GLOSSARY

ADA	Americans with Disabilities Act
ADMS	Archived Data Management Systems
ADUS	Archive Data Management User Service
AFC	Automated Fare Collection
APC	Automated Passenger Counting
API	Application Programming Interface
APTA	American Public Transportation Association
APTS	Advanced Public Transportation Systems
ASP	Active Server Pages (product of Microsoft)
ASTM	American Society for Testing and Materials
ATIS	Advanced Traveler Information System
AVL	Automated Vehicle Location
BSI	Bus Stop Inventory. This term may apply to the physical database inventory or to the TSC Bus Stop Inventory Best Practices Guide.
CAD	Computer Aided Dispatch
COM	Component Object Model
CORBA	Common Object Request Broker Architecture
COTS	Commercial Off the Shelf [software]
CSR	Customer Service Representative
DBA	Database Administrator
DBMS	Database Management System
DOT	Department of Transportation
DTD	Document Type Definition
FGDC	Federal Geographic Data Committee
FTA	Federal Transit Administration
FTE	Full Time Equivalent
GDF	Geographic Data File
GIS	Geographic Information System
GML	Geographic Markup Language
GOS	Geospatial One-Stop
GPS	Global Positioning System
GUI	Graphical User Interface
HOV	High Occupancy Vehicle
HTML	Hyper Text Markup Language
IEEE	Institute of Electrical and Electronics Engineers
IP	Intellectual Property
ISO	International Standards Organization
IT	Information Technology
ITE	Institute of Transportation Engineers
ITS	Intelligent Transportation System
IVR	Interactive Voice Response
LRG	Location Referencing Guidebook
LRM	Location Referencing Method
LRS	Location Referencing System
MPO	Metropolitan Planning Organization
MOM	Message Oriented Middleware

NSDI	National Spatial Data Infrastructure
NTCIP	National Transportation Communications for ITS Protocols
OSS	Open Source Software
RDBMS	Relational Database Management Systems
RFP	Request for Proposals
RMI	Remote Method Invocation
RT	Real Time
SAE	Society of Automotive Engineers
SMMS	Spatial Metadata Management System (product of Intergraph)
SOAP	Simple Object Access Protocol
SQL	Structured Query Language
TAZ	Transportation Analysis Zone
TCIP	Transit Communication Interface Profile
TCIP2	TCIP Phase 2 refers to the effort to develop “Dialogs”
TCRP	Transit Cooperative Research Program
TrMS	Transit Management Subsystem
TrVS	Transit Vehicle Subsystem
TSC	Transit Standards Consortium
TSP	Transit Signal Priority
TWG	Technical Working Group (of the TCIP Standards Committee)
UML	Unified Modeling Language
USDOT	United States Department of Transportation
XML	Extensible Markup Language
XSLT	Extensible Stylesheet Language Transformations
WFS	Web Feature Service

## **GLOSSARY**

<b>TERM</b>	<b>DEFINITION</b>	<b>SOURCE</b>
accuracy	closeness of agreement between a test result and the accepted reference value (NOTE A test result can be observations or measurements.)	ISO 19113
altitude	Elevation above or below a reference datum, as defined in ANSI X3.61 (ANSI, 1986); the z-value in a spatial address.	SAE J1746
application schema	conceptual schema for data required by one or more applications	ISO 15046
authority table	A table or file that serves as the definitive source for data values in an organization	
base map	The foundational spatial and attribute data to which transit data are referenced	Section 5 of these guidelines
centerline file	A database of arcs and nodes describing network geometry (shape) but not necessarily network connectivity, thus a centerline file may or may not be navigable.	Section 5 of these guidelines
conceptual model	model that defines concepts of a universe of discourse	ISO 15046

TERM	DEFINITION	SOURCE
conceptual schema	schema of a conceptual model	ISO 15046
conversion	change of coordinates, based on a one-to-one relationship, from one coordinate system to another based on the same datum (An example is between geodetic and Cartesian coordinate systems. A conversion uses parameters which are constant.)	ISO 15046
Coordinated Universal Time:	UTC time scale maintained by the International Bureau of Weights and Measures and the International Earth Rotation Service (IERS) that forms the basis of a coordinated dissemination of standard frequencies and time signals	ISO 15046
coordinates	Pairs of numbers expressing horizontal distances along orthogonal axes; alternatively, triplets of numbers measuring horizontal and vertical distances.	SAE J1746
cycle	regular recurrence of an event or a state	ISO 19108
data element	A formalized representation of some information (such as a fact, a proposition, an observation, etc.) about some concrete or abstract thing—an entity—such as a person, place, process, object, concept, association, state, or event.	SAE J1746
datum	A set of parameters and control points used to accurately define the three-dimensional shape of the Earth (e.g. as an ellipsoid). The corresponding datum is the basis for a planar coordinate system.	SAE J1746
duration	quantity of time (duration is the temporal equivalent of length)	ISO 19108
edge	1-dimensional topological primitive	ISO 19108
entity	Anything of interest (such as a person, place, process, property, object, concept, association, state or event) within a given domain of discourse (in this case, within the ITS domain of discourse).	IEEE 1489-1999
Event [transit]	[transit] action which occurs at a time, location or both	adapted from NTCIP 1404:2000
Feature	Points, lines or polygons with associated attribute information.	
feature catalogue	definition and description of the features, together with their feature functions, feature attributes, and feature relationships, occurring in one or more sets of geographic data.	ISO 15046
feature fusion	feature succession in which two or more previously existing instances of a feature type are replaced by a single instance of the same feature type	ISO 19108
feature substitution	feature succession in which one feature instance is replaced by another feature instance of the same or different feature type	ISO 19108
feature succession	replacement of one or more feature instances by other feature instances, such that the first feature	ISO 19108



TERM	DEFINITION	SOURCE
	instances cease to exist	
gazetteer	directory of a class of real-world phenomena containing information regarding position	ISO 15046
Geodetic Datum	A mathematical model of the Earth's shape. A geometric set of five quantities which serves as a locational reference or base for other quantities. The five quantities are the latitude and longitude of an initial point, the azimuth of a line from this point and two constants necessary to define the terrestrial spheroid.	SAE J1746
Geographic Information System (GIS)	GIS is a System composed of computer software, hardware and data, and personnel to help manipulate, analyze and present information that is tied to a spatial location	ESRI: What is GIS?
Gregorian calendar	calendar in general use introduced in 1582 to correct an error in the Julian calendar (In the Gregorian calendar common years have 365 days and leap years 366 days divided into 12 sequential months.)	ISO 19108
Information Service Provider (ISP)	A public or private entity which may be responsible for gathering, fusing, analyzing, and/or reporting transportation related information to users, including vehicles and non-mobile users.	SAE J1746
instant	point representing position in time	ISO 19108
intersection	The geometrical crossing or touching of two or more linear transportation features such that some points in space are shared between the features. Features may be paths, roads, railroads, waterways, or air routes. Intersections may be physical and topological (for example an at-grade highway-rail intersection) or may be logical and non-topological (for example an overpass of one road over another).	SAE J1746
interval scale	A scale with an arbitrary origin that can be used to describe both ordering of values and distances between values.	ISO 19108
Level of service	A measurable level of service provided such as Transit service performance indicators or vendor support service standards for providing IT support	
linear referencing method	A mechanism for finding and stating the location of an unknown point along a network by referencing it to a known point.	NCHRP 20-27(2)
linear referencing system	The policies, records, office and field procedures associated with a set of linear referencing methods along with a means for transforming among various methods.	NCHRP 20-27(2)
link	A topological connection between two nodes. A link may contain additional intermediate coordinates (shape points) to better represent the	SAE J1746

TERM	DEFINITION	SOURCE
	shape of curved features. A link may be directed by ordering its nodes.	
life span	period during which something exists	ISO 19108
Location Referencing Method	System of determining the position of an entity relative to other entities or to some external frame of reference.	Adapted from SAE J1746
metadata	A self-description about the content, quality, condition and other characteristics of a single piece of data or entire file of data.	ISO 19115
metadata element	discrete unit of metadata (Equivalent to an attribute in UML terminology. Metadata elements are unique within a metadata entity.)	ISO 19115
metadata entity	set of metadata elements describing the same aspect of data (May contain one or more metadata entities. Equivalent to a class in UML terminology.)	ISO 19115
metadata management	Creation, organization, storage, control, access and query of metadata. (Metadata management is to metadata what database management is to data.)	
NAD (North American Datum)	The official reference ellipsoid used for the primary geodetic network in North America.	SAE J1746
NAD27 (North American Datum 1927)	Geodetic datum defined by the geographic position of triangulation station Meades Ranch and the azimuth from that station to station Waldo on the Clarke spheroid of 1866.	SAE J1746
NAD83 (North American Datum 1983)	Geodetic datum defining an Earth-fitting ellipsoid and for relating existing coordinate systems to the Earth's center of mass. It is based on both satellite and terrestrial data used in developing the Geodetic Reference System of 1980 and later the World Geodetic System of 1984 (WGS-84). This improved Earth-centered model slightly changes the latitude and longitude of almost every point in North America with respect to the earlier NAD27 datum.	SAE J1746
National ITS Architecture	A framework established by the USDOT, for planning, defining, and integrating intelligent transportation systems	
node	0-dimensional topological primitive	ISO 19108
ordinal scale	scale which provides a basis measuring only the relative position of an object	ISO 19108
quality	totality of characteristics of a product that bear on its ability to satisfy stated and implied needs	ISO 19113
schema	formal description of a model	ISO 19107
schemata	plural of schema	
state	condition that persists for a period	ISO 19108
temporal coordinate system	temporal reference system based on an interval scale on which distance is measured as a multiple	ISO 19108

TERM	DEFINITION	SOURCE
	of a single unit of time	
temporal position	location relative to a temporal reference system	ISO 19108
temporal reference system	Reference system against which time is measured	ISO 19108
topology	The logical relationships among map features in a digital base map. It can be used to characterize spatial relationships such as connectivity and adjacency.	SAE J1746
transformation	change of coordinates, based on a one-to-one relationship, from one coordinate reference system to another based on a different datum. (A transformation uses parameters which may have to be derived empirically by a set of points common to both coordinate reference systems).	ISO 15046
transit feature	Real-world transit objects or events which are represented in a GIS, as points, lines, or polygons with associated attribute information.	
transit network	A subset of a transportation network that describes the public transportation characteristics of the base map.	
transportation network	A two-dimensional graph of links (or arcs) and nodes, describing network connectivity and not necessarily network geometry (shape) but provides the basis for analytical operations such as pathfinding and flow. See also Base Map.	NCHRP 20-27(2)
World Geodetic System of 1984 (WGS-84)	An earth-centered global reference frame, including an earth model, based on satellite and terrestrial data. It contains primary parameters that define the shape, angular velocity, and the earth mass of an earth ellipsoid, and secondary parameters that define a gravity model of the earth. Primary parameters are used to derive latitude-longitude coordinates (horizontal datum). Secondary parameters are used for determining the orbits of GPS navigation satellites and define a coarse vertical datum; therefore WGS-84 is a complete geodetic system. EGM-96 is a refined model established in 1996.	SAE J1746